Aspects of the Development of a 1–5 μm Infrared Camera for Astronomy

John Thornton Rayner

Presented for the Degree of Doctor of Philosophy

University of Edinburgh

June 1988

This thesis has been composed by myself and consists entirely of my own work, except where noted in the text.

June 1988

To my Mother and Father

Contents

Abstract

Aspects of the design, development and evaluation of a 1–5 μm near–infrared camera for a large ground–based telescope are described and discussed. In particular this thesis will consider the detailed cryogenic design and the closely inter–related areas of the optical and mechanical design. Also discussed is near–infrared imaging and imaging polarimetry of the star forming region OMC 2, acquired during the astronomical commissioning of the camera.

Acknowledgements

Undoubtedly the two people I have most to thank are my supervisor, Ian McLean, and Tim Chuter, respectively IRCAM Project Scientist and Project Engineer. This thesis would have been *impossible* without them and I feel very fortunate indeed to have benefited from their advice and guidance.

It is a great pleasure for me to thank the whole of the IRCAM project team and in particular Ron Beetles, John Clark and John Harris for teaching me a lot about the realities of engineering design, and for never once refusing a request for help. When they weren't debating how much whisky they could transfer through Customs, Bill Duncan and Jim Radostitz did sometimes get around to discussing cryogenics with me, for which I'm very grateful. I'm not so grateful for the exploding noises 'Big Foot' Duncan used to make whenever I attempted helium transfers. I received an unforgettable workshop course from Frank Holmes, during which Frank demonstrated how tenuous our hold on life really is. The help and advice received from Frank and the workshop staff is very much appreciated. Thanks also to Terry Lee and Malcolm Smith for lying on my behalf, to Donald Pettie for sharing with me his always interesting views on engineering the universe and how it all can be documented in one page, and to John Cooke for his encouragement. Hans Zinnecker was always more than willing to discuss the astronomy with me, but more than made up for it by hardly ever passing the ball. Thankyou to Marjorie Fretwell for providing the drawings, to the staff of Photolabs, Bill, Duncan, Eric and Brian, to the Library staff, Angus, Anne *et al.* , and to Liz Gibson for providing a shoulder to cry on (Stan Blackley lent me the step ladders). I suppose I should thank my fellow students, for an assortment of football injuries, James Dunlop (cracked rib), Andrew Mead (a cold, caught while waiting in vain for a decent pass), James More (hæmatoma of the right quadrucep), and Mark Yates for keeping me well informed (top draw, Malcolm's office, ha, ha, ha).

Colin Aspin went above and beyond the call of duty by giving me a place to stay during my 5 months in Hawaii, while commissioning the camera (35mm). His help with the data reduction has been invaluable. Last and least (they understand) thankyou to my fellow outcasts, Mark Bird, for letting me have the office swivel chair, and to Mark McCaughrean for sometimes shutting up. This work was funded by SERC through a Research Studentship.

Chapter 1

Introduction

Astronomy in the near–infrared wavelength range from 1–5 μm offers unique insights into such important and diverse topics as star formation and cosmology. At birth stars are so embedded in obscuring clouds of gas and dust, that optical imaging observations with CCDs cannot reveal their stellar content. Infrared radiation penetrates these regions far more effectively than optical radiation. For example, at a wavelength of only 2.2 μm (four times mid–visible), the attenuation is up to 10^{10} less (equivalent to a visible extinction of A_V =25 magnitudes). Thus the near–infrared is a critical portion of the spectrum for the study of star formation. However, until recently, observations in this wavelength range have been restricted to mapping with single element photometers by raster scanning the telescope, a laborious and inefficient technique. With the advent of high quality two–dimensional near–infrared detector arrays, this restriction has been removed. "Common–user" access to true infrared cameras and infrared spectrometers based on detector arrays, promises to revolutionise the subject.

Infrared sources in star formation regions are often extended and associated with considerable nebulosity, regions where an *imaging* capability is essential to disentangle the complex morphologies. Infrared detector arrays will allow for the first time deep surveys of the sky at infrared wavelengths. For galaxies at redshifts of z$\sim$1, the familiar optical starlight is redshifted into the near–infrared, while for objects, such as primæval protogalaxies at z= 10, even Lyα emission (λ = 0.1216 μm) would be redshifted into the near–infrared. The study of galactic evolution at high redshifts and the search for primæval galaxies stand out as the main potential rewards of a deep infrared survey.

Clearly near–infrared detector arrays offer the opportunity to tackle some of the most important problems facing modern astrophysics.

Foreseeing the commercial availability of high quality detector arrays to the astronomical community, a new generation of infrared instruments was conceived at the Royal Observatory, Edinburgh (ROE), which would employ arrays in cameras and spectrometers. These instruments were to be common–user, engineered to the highest standards, reliable and easy to use. In 1981 a project was started to build an Imaging Spectro–Polarimeter (ISP) based on an optical array CCD. This initiative not only made possible significant astronomical observations using the ISP during 1983 and 1984 but also succeeded in achieving the primary objective of gaining expertise in detector arrays needed to advance the use of infrared arrays in astronomy. Several years of planning, design and engineering by the ROE, on behalf of the UK's infrared astronomers led in September 1986 to the delivery to the UK Infrared Telescope (UKIRT) on Mauna Kea, Hawaii, of the first of the new generation of infrared cameras (IRCAM), employing a detector array sensitive in the near–infrared from 1–5 μm.

A critical part of the specification for the new camera was the vacuum–cryogenic design and performance, for which the author was responsible, and which in large part forms the subject of this thesis. As will become clear, this work has of necessity included considerable attention to the inter–related areas of optical and mechanical design. The remainder of the thesis concerns the astronomical application of the cryogenic camera to a region of star formation.

In a very real sense it is not possible to consider any part of a system, in this case an infrared camera, in isolation. This includes the impact of time and cost considerations as well as the engineering design. Ultimately however, the design of IRCAM has been driven by the need to embody several astronomically desirable configurations into a single instrument: imaging at different image scales, photometric imaging through a selection of broad and narrow band filters, imaging polarimetry, but principally, high spectral resolution imaging using Fabry–Pérot étalons. The most important physical constraint on the design of the camera concerns optically matching the fixed pixel size of the detector array to the 3.8m telescope. The requirements of high spatial resolution and high spectral resolution, mean that the detectors used in IRCAM must operate under low background conditions. To achieve (sky and telescope) background limited

performance under these conditions, the array must be cooled to a temperature between liquid helium (4K) and solid nitrogen (50K) in order to minimise thermally generated dark current in the detector, and as much of the optics as practicable must be cooled to liquid nitrogen temperatures (77K) to minimise thermal background photons falling onto the array. Chapter 2 examines the size of the unwanted backgrounds falling onto the array, together with the extent to which cryogenic cooling can be used to reduce them, the nature of dark current in the detector and the degree of cooling required to eliminate it, and the choice of a liquid nitrogen/liquid helium dewar to cool the camera. As will be shown in Chapter 3, the decision to fully exploit high spectral resolution imaging while at the same time making the most of UKIRT's unvignetted field of view, has had a major impact on the size and layout of the optical components. In Chapter 3, optically interfacing the detector array with a large telescope will be discussed, along with the precautions taken to baffle the camera against stray flux which would otherwise significantly degrade performance. The detailed cryogenic design of IRCAM is the subject of Chapter 4. This has entailed developing support structures which have high stiffness and low thermal conductivity but which also accommodate the relative thermal contraction between the cryogenic and ambient environments. A thermal model of the cryostat was developed and successfully used to optimise the design, which includes the cryogenic mechanisms comprising the dewar/camera head interface, the detector and focus assembly, and the filter wheels. By visiting other establishments during the design phase, it became clear that a great deal of the design of astronomical cryogenic mechanisms was done empirically , and with little attempt at modelling. The cryogenic mechanisms used in IRCAM have been extensively modelled to ensure that they meet the carefully considered design goals.

In 1984, the UKIRT Unit (on behalf of the ROE) placed an order with the Santa Barbara Research Center (SBRC) for several infrared array detectors. As specified by SBRC (Orias *et al.* 1986, see Table 1.1), the performance of the device was very good indeed. It was known that considerable improvements in dark current would result when the device was operated colder than 50K (see Chapter 2), but that the performance of the silicon multiplexer became poor in the region 20–30K. Since other detector properties are also a function of temperature, it is vital for the operating temperature of the detector to be kept very stable. For these reasons the camera design specified that the array should be capable of operating at any selectable and stable ($\pm$0.1K)

Table 1.1: **InSb Array Specifications** (high doped device)

Detector material	InSb
Format	62x58 (3596 pixels)
Pixel spacing	76 μm x 76 μm
Active area fill factor	90%
Readout noise	~380 electrons
Full well capacity	$\sim 10^6$ electrons
Working pixels	92%
Quantum efficiency	>60% over 1–5 μm
	>80% at 2 μm
Dark current at 40mV	<15x10^4 electrons s^{-1}
reverse bias, 50K	

temperature in the range 10–50K. A new device called a Variable Conductance Gas Switch was successfully developed specifically for this purpose and forms the subject of Chapter 5. (In modified form this chapter has already been published as a paper (Rayner *et al.* 1988).)

Following the delivery of IRCAM to UKIRT, the author participated in the scientific commissioning of the camera. One of the regions observed during the commissioning phase was Orion Molecular Cloud 2, a star formation region located about 12$\hat{'}$ north of the Trapezium cluster in M42. These observations and their interpretation are considered in Chapter 6. This region was chosen for study not just for its intrinsic interest as a region of intermediate–mass star formation to contrast with the high–mass star formation region of the Trapezium OB cluster, but also because it had been mapped extensively with single channel photometers at several infrared wavelengths, providing us with calibration data during commissioning of the new instrument. The wide spatial coverage and high spatial resolution, together with the imaging polarimetry capabilities offered by IRCAM, have added significantly to our understanding of OMC 2.

As such, the cryogenic camera, IRCAM, is just one component of the infrared imaging system. Before discussing the camera in detail, we will briefly describe the system it is built into (see Figure 1.1).

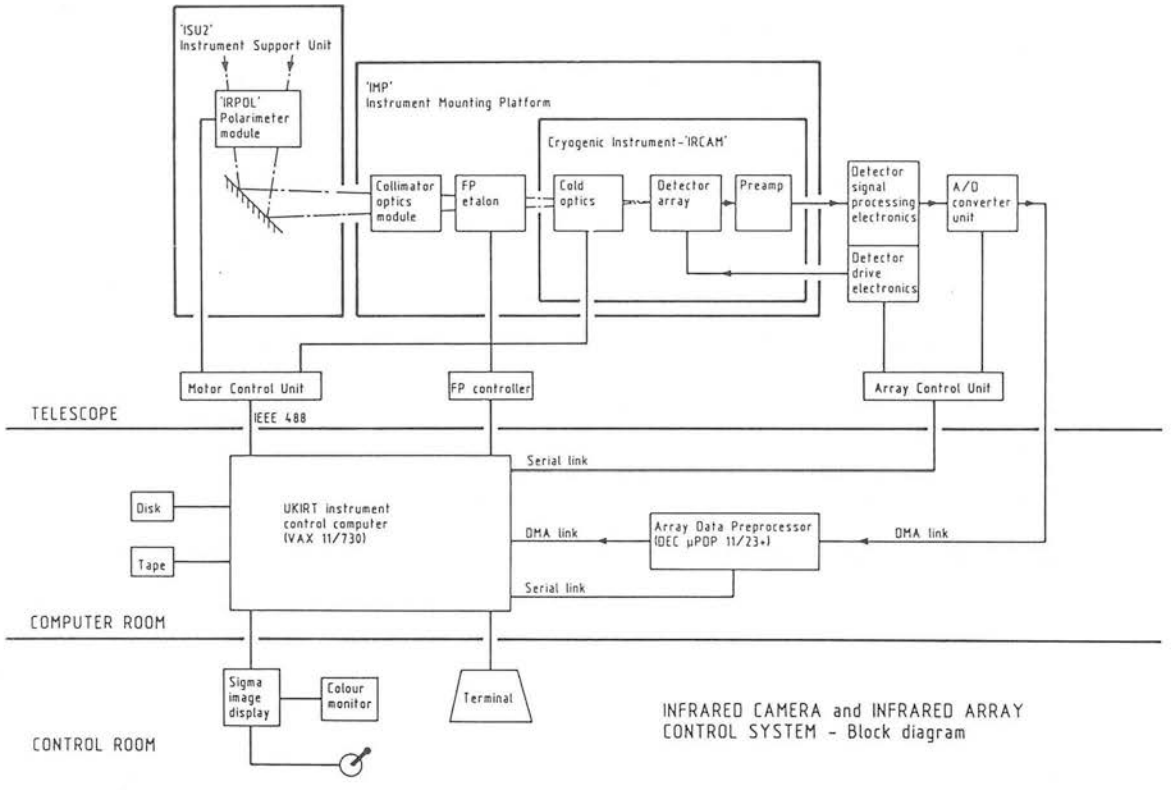


Figure 1.1 : Infrared camera and infrared array control system.

Light emerging from the UKIRT Instrument Support Unit (ISU2) comes to focus in front of the collimating optics. The diverging f/36 beam is collimated and redirected into the IRCAM Camera Head by gold coated off–axis mirrors, where it is refocused by the cold optics onto the detector array. By selection of an appropriate lens prior to cool–down of the cryostat, IRCAM can be set to any one of three image scales, namely 0.6, 1.2 or 2.4 $\widehat{''}$pixel^{-1}. Both the collimator and cryostat are seated on a large Instrument Mounting Platform (IMP) protruding from ISU2 and braced with struts on the telescope mirror cell. A Fabry–Pérot étalon can be placed in the collimated beam in front of the cryostat entrance window, and an infrared polarimeter module (IRPOL) can be remotely lowered into the beam just above the dichroic beam–splitter in ISU2. Crucial drive electronics for the detector array together with some signal processing hardware and the A/D converter for the output signal, are located under the IMP, as close as possible to the cryostat. In a cabinet attached to the mirror cell is the Array Timing Generator/Computer Interface and the Array Control electronics. This unit decodes the high–level commands from the UKIRT VAX computer into procedures recognisable by the electronic drive system and synchronises the digitisation of the output from the array. Also mounted on the mirror cell is the motor control unit required to drive the filter wheels and detector focus. Data from the array is digitised and sent to a remote pre–processing computer where many rapidly accumulated exposures can be coadded into a single image before transfer to the UKIRT VAX 11/730 computer for final storage, analysis and display.

Chapter 2

The Need for Cryogenics

2.1 Introduction

There are two fundamental reasons why infrared cameras need to be cooled to low temperatures :

- Thermally–generated dark current from the detector array must be reduced to a level where it is negligible compared to the observed sky background, otherwise the sensitivity of the camera system is drastically impaired
- Unwanted thermal background emission seen by the detector must be reduced to a level where it is negligible compared to the observed sky background or detector array dark current, whichever is greater

In an intrinsic photovoltaic semiconductor, such as indium antimonide (InSb), the photosignal is produced by the excitation of electron–hole pairs by photons having energy at least as great as the band–gap energy, E_g :

$$\frac{hc}{\lambda} \geq E_g \tag{2.1}$$

where h is the Planck constant, c the velocity of light and λ the photon wavelength. The long wavelength limit λ_c, beyond which a photon has insufficient energy to generate an electron–hole pair, is given by :

$$\lambda_c = \frac{hc}{E_g} \tag{2.2}$$

Electron–hole pairs are also produced by thermal processes in the semiconductor material. InSb is sensitive to a long wavelength limit of about 5 μm, having a relatively small band–gap (E_g=0.24eV) compared to silicon (E_g=1.14eV, λ_c = 1.1 μm). Optically sensitive astronomical Si CCDs need to be cooled to temperatures of ~140K for their thermally generated signal or "dark current" to be negligible compared to the photosignal. Because of the smaller band–gap and correspondingly greater dark current, InSb is cooled to much lower temperatures.

As a consequence of cooling to remove the dark current, a detector becomes more sensitive to all other sources of photosignal, including unwanted background signals. A typical configuration of an astronomical imaging system is that of a warm ground–based telescope forming an image of the sky in its focal plane, followed by a camera containing spectral bandwidth limiting filters and optics that re–image the telescope focal plane onto the detector array. The background signals are dependent upon the wavelength of observation, and at wavelengths beyond about 2 μm the major source of background is thermal emission from warm objects in the detectors field of view. These include the atmosphere above the telescope, the telescope itself, and the components comprising the camera. Of course the ideal solution is to cool the entire telescope and camera, and place it above the atmosphere. Actively cooling a large ground–based telescope is not practicable, and the best that can be achieved is to place it on a high, cold site, and to minimise its effective emissivity. However, actively cooling the camera and its detector is by comparison relatively easy.

In this chapter we will investigate the magnitude of the thermal dark current in the detector and the thermal backgrounds seen by the detector, and therefore predict the degree of cooling required to reduce these two unwanted signals to optimum levels for ground–based near–IR imaging. The choice of cooling system is discussed and the final choice for cooling IRCAM is described. First, we briefly examine how the signal is detected in the array.

2.2 Signal Detection

The infrared detector array used in IRCAM was built by Santa Barbara Research Center (SBRC) and consists of a 62 x 58 array of photo–voltaic InSb photodiodes

bump–bonded to a silicon substrate that contains the MOSFET direct read–out (DRO) multiplexer. InSb is a soft semi–metal and is easy to work with compared to other potential detector materials (such as Cadmium–Mercury–Telluride, CMT); a fact which has contributed to its mature processing technology. The detectors (see Figure 2.1) are fabricated on a 500 μm thick n–type substrate which is grown in bulk phase with a donor concentration of $N_D \sim 10^{18}$ cm^{-3}. The p–type regions are ion implanted with an acceptor concentration of $N_A \sim 10^{15}$ cm^{-3}. Following passivation (insulation with SiO_2) of the interjunction regions, metal gates are fabricated between the junctions. These gates control the surface potential between the p–n diodes for reducing dark current. Insulating cover–glass is then applied over the gates. Indium 'bumps' of about 15 μm diameter are fabricated on the p–type regions of the detector array, which is mated (hybridised) with the similarly bump–bonded multiplexer by accurately aligning the two chips and then applying pressure to cold weld the detector and multiplexer indium bump. The space remaining between the InSb array and Si multiplexer is filled with an epoxy resin. Although the hybrid chip produced by bump–bonding can survive cryogenic cycling, the epoxy is needed for the mechanical stress of thinning the n–type substrate from 500 μmto $\sim$20 μm. To avoid scattering due to the metal gates in the multiplexer, the arrays are illuminated from the n–type side (backside–illumination). If the n–type substrate is too thick a photo–generated hole (a minority carrier in the n-type material) will recombine before it can diffuse into the junction region where recombination is prevented by the junction electric field. (In a detector under reverse bias, minority carriers arriving near the depletion region are swept across the junction by the electric field there, into the p–type material where they become majority carriers, generating a reverse current which discharges the junction.) On the other hand if the substrate is too thin IR photons will not be efficiently absorbed. (Thinning still damages a significant number of arrays and, as a result, SBRC are currently considering epitaxial growth of InSb wafers to the required 20 μm.)

The arrays of InSb detectors are read out using a direct integration source follower for each detector circuit (see Figure 2.2 and Orias *et al.* 1986). Detector current (I_{det}) is integrated on the detector capacitance (C_{det}) and the input capacitance of the multiplexer read–out (C_{mos}). An integration is started by opening the reset MOSFET switch. This allows the reverse bias voltage on the detector pn junction,$V_d = V_{rst} - V_{sub}$ to discharge due to I_{det}. At the end of an integration the voltage source follower MOS-

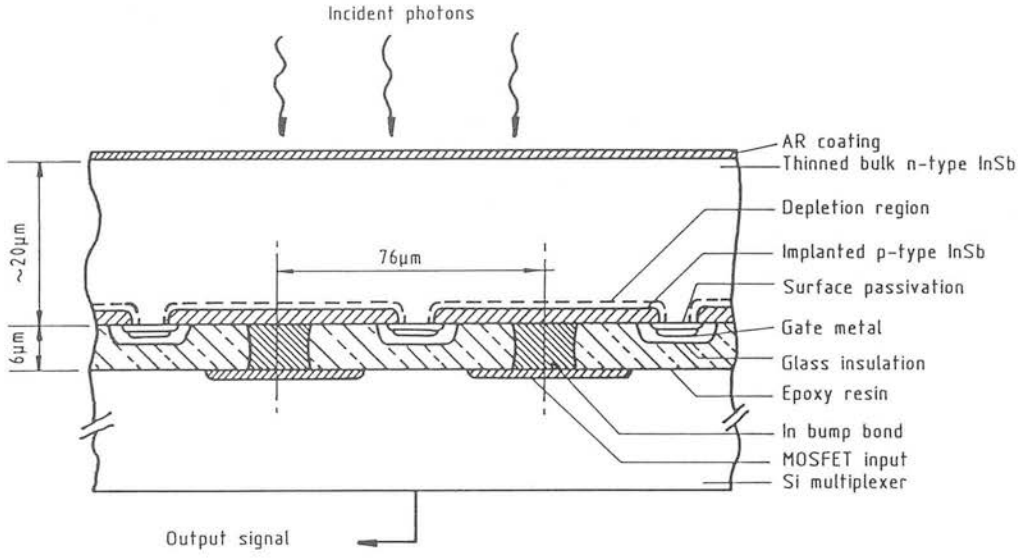


Figure 2.1 : Physical layout of the InSb array. The detector array consists of an n–type substrate with p–type regions implanted within it. Indium bumps are fabricated on the p–type regions of the detector array, which is mated with a similarly bump–bonded Si multiplexer by applying pressure.

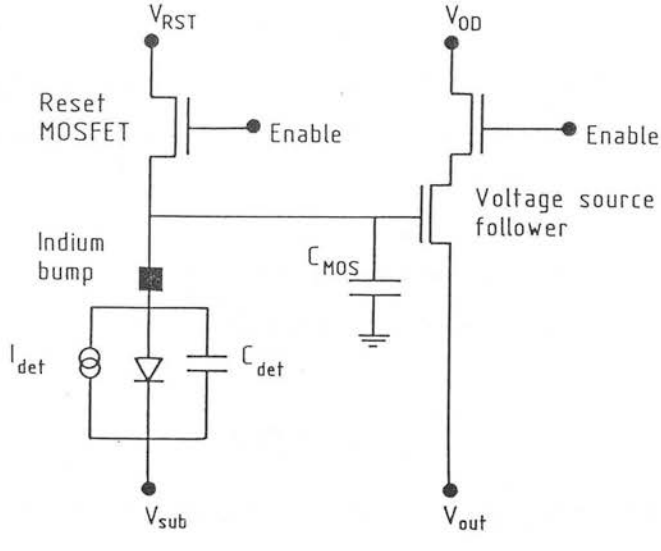


Figure 2.2 : Read–out circuit for the InSb array. Each InSb detector is read out using a direct integration source follower for each detector circuit. The Si read–out device is a CRC–228 Direct Read–out (DRO), an n–channel MOSFET switching device.

FET is enabled and the integrated voltage is (non–destructively) readout. The difference in the output voltage before and after the reset pulse ($V_o(t) - V_o(0)$) is a function of the discharge produced by incident IR photons (I_{photon}) and detector dark current (I_{dark}). By definition the capacitance of a device C, is the ratio of the charge Q stored on it to the voltage V across it, from which it follows that (since C is a function of V) :

$$dQ = \left(C + V\frac{\partial C}{\partial V}\right) dV \tag{2.3}$$

Also by definition :

$$dQ = I_{det}dt = (I_{photon} + I_{dark})dt \tag{2.4}$$

Combining Equations 2.3 and 2.4 gives :

$$\frac{dV}{dt} = \frac{I_{photon} + I_{dark}}{C + V\frac{\partial C}{\partial V}} \tag{2.5}$$

Therefore the voltage integrated onto the detector is given by :

$$V_o(0) - V_o(t) = A_s \int_0^t \frac{I_{photon} + I_{dark}}{C + V_d\frac{\partial C}{\partial V}}\, dt \tag{2.6}$$

where the total capacitance $C = C_{det} + C_{mos}$, and A_s is the voltage source follower gain ($\simeq$0.8).

A disadvantage of the voltage source follower design is that the detector bias V is not maintained at zero volts. Vampola (1982) shows that I_{dark} and C_{det} are functions of the detector bias with the result that signal detection (Equation 2.6) is inherently non-linear. Modelling of Vampola's equations by McCaughrean (1988) has shown that the non-linearity is at most 10% and McCaughrean also discusses algorithms to correct for the non-linearity.

The number of electrons required to fully discharge the initial reverse bias on a detector is known as the "well depth" and can be thought of as the integrating storage capacity :

$$\text{Well depth} = \frac{C_{det}V_i}{q} \tag{2.7}$$

where C_{det} is the detector capacitance (typically 0.7pF), V_i is the initial reverse bias (typically -250mV) and q is the charge on an electron, resulting in a well depth of $\sim 10^6$ electrons. Although the detector capacitance changes slightly as it discharges, this figure is very useful to keep in mind when comparing competing signal currents.

To measure the photo–current (and thence flux) from an astronomical source we need to take at least two equal length integrations of time t, one of the source plus background (sky plus telescope, instrument and dark current), and one of the background alone :

$$\text{Image 1} \; : \; (I_{source} + I_{sky} + I_{instr} + I_{dark})t$$

$$\text{Image 2} \; : \; (I_{sky} + I_{instr} + I_{dark})t$$

Assuming linearity (linear superposition) :

$$I_{source} = (I_{source} + I_{sky} + I_{instr} + I_{dark}) - (I_{sky} + I_{instr} + I_{dark}) \tag{2.8}$$

If the read noise is assumed to be independent of the integration time, whereas the other current sources obey Poisson statistics it follows that the signal to noise ratio (S/N) per detector of the source photo–current is given :

$$\frac{I_{source}t}{[t(I_{source} + 2I_{sky} + 2I_{instr} + 2I_{dark}) + 2N_R^2]^{\frac{1}{2}}} \tag{2.9}$$

where I_{source} is the source photocurrent (electrons s^{-1})
I_{sky} is the photo–current due to the sky and telescope background (electrons s^{-1})
I_{instr} is the photo–current due to the instrument's thermal background (electrons s^{-1})
I_{dark} is the detector dark current (electrons s^{-1})
N_R is the detector read noise (electrons)

It is important to appreciate that only the assumption of linearity allows us to solve for I_{source} by a simple subtraction (linear superposition). Equation 2.9 is a quantitative statement of the obvious – to improve measurement of the source signal, unwanted background signals must be minimised. In the infrared the source photocurrent is usually very small compared to some of the other current sources in Equation 2.9. When $I_{dark} \ll I_{sky}$, $I_{instr} \ll I_{sky}$ and $N_R^2 \ll I_{sky}t$, the detector is said to be sky or background limited in performance (BLIP). A good detector must have $N_R^2 \ll$ well depth. For the SBRC InSb array used in IRCAM $N_R \simeq 500$ electrons RMS and the well depth $\simeq 10^6$.) In the following sections we will investigate the nature of the dark current and background photocurrents, and ways of reducing them.

2.3 Dark Current

Dark current arises from thermal generation of electron–hole pairs in the bulk and depletion regions of the InSb, and from other surface and leakage currents in the detector and DRO. Diffusion current is due to thermal generation and recombination of carriers outside the depletion region. In the absence of a significant electric field, carriers reach the boundary of the depletion region by diffusion. Under reverse bias, minority carriers (holes) which have been generated near or have diffused to the edge of the depletion region, are swept across the junction under the influence of the electric field there, resulting in a reverse current. Generation–recombination current (G–R) is due to the thermal generation and recombination of carriers within the depletion region. For temperatures below about 110K, diffusion current is negligible compared with G–R current . G–R current is described by the expression (Vampola 1982) :

$$I_{g-r} = \frac{2kTA_d}{q(R_oA_d)_{g-r}}(1 - V_d/V_{bi})^{\frac{1}{2}}[exp(qV_d/kT) - 1] \tag{2.10}$$

where A_d is the detector area

k is the Boltzmann constant

T is the absolute temperature

q is the charge on an electron

V_{bi}, V_d, the built in junction potential, and the applied bias voltage

R_o is the dynamic detector resistance at zero bias (V_d=0)

and $(R_oA_d)_{g-r}$ its normalised figure-of-merit.

The dynamic detector resistance at zero bias R_o is found by taking the derivative of the detector current with respect to voltage and as such the normalised R_oA product is the most commonly measured and intercompared parameter for photovoltaic detectors. The R_oA product when controlled by G-R current is given by (Vampola 1982) :

$$(R_oA_d)_{g-r} = \frac{4kT\tau_o}{q^2n_iV_{bi}^{\frac{1}{2}}}\left[\frac{2\varepsilon_s}{q}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)\right]^{-\frac{1}{2}} \tag{2.11}$$

where τ_o is the effective minority carrier lifetime in the depletion region

N_A is the ionised acceptor concentration on the p side of the junction

N_D is the ionised donor concentration on the n side of the junction

ε_s is the dielectric permittivity for InSb

n_i is the intrinsic carrier concentration

The intrinsic carrier concentration can be calculated from (Yang 1978) :

$$n_i^2 = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{\frac{3}{2}} 2\left(\frac{2\pi m_p^* kT}{h^2}\right)^{\frac{3}{2}} \exp(-E_g/kT) \tag{2.12}$$

where m_n^* is the effective mass of the minority carriers (holes) on the n side
m_p^* is the effective mass of the minority carriers (electrons) on the p side of the junction
h is the Planck constant
E_g is the InSb band–gap

For InSb, $m_n^* = 0.015m_e$, $m_p^* = 0.39m_e$, and $E_g \simeq 0.24 - 2\text{x}10^{-4}T$ (Vampola 1982) :

$$n_i = 3.26\text{x}10^{14}T^{\frac{3}{2}} \exp(-1400/T) \quad [\text{cm}^{-3}] \tag{2.13}$$

The theoretical dependence of G–R dominated dark current on temperature has been calculated using Equations 2.10, 2.11 and 2.12 for high doped ($N_A = 10^{18}$ cm^{-3}, N_D=2x10^{15} cm^{-3}) and low doped ($N_A = 10^{18}$ cm^{-3}, N_D=4x10^{14} cm^{-3}) detectors. Measurements of $(\text{R}_o\text{A}_d)_{g-r}$ have been used to determine an effective minority carrier lifetime of τ_o=2x10^{-8}sec (Thom and Yang 1986), the other parameters are from Vampola (1982). In Figure 2.3 G–R current (I_{g-r}) is plotted against temperature together with data (scaled from R_oA_d products) from a number of high–doped SBRC PV InSb arrays (A_d=75 μm^2 and V_d=–40mV, Thom and Yang 1986). Although E_g, V_{bi} and τ_o are temperature dependent, the effects are small compared to the temperature dependence of n_i. At temperatures easily produced by pumping over liquid nitrogen ($\simeq$47K), the G-R current falls to a level where narrow band imaging ($\lambda/\Delta\lambda$ $\sim$100, see Figure 2.6) becomes possible. However, prior to manufacture of the SBRC 62x58 InSb+DRO arrays, the available data (see Figure 2.3) showed only that measured dark currents followed the predicted exponential decrease until temperatures of $\sim$55K were reached. Below 55K new leakage sources, showing only a weak temperature dependence, became significant. These leakage sources, discussed by Thom and Yang (1986), include :

- Tunneling due to the junction electric field
- Surface leakage (controlled by V_{gate})
- Leakage through the gate oxide
- Photon leaks associated with the cryostat

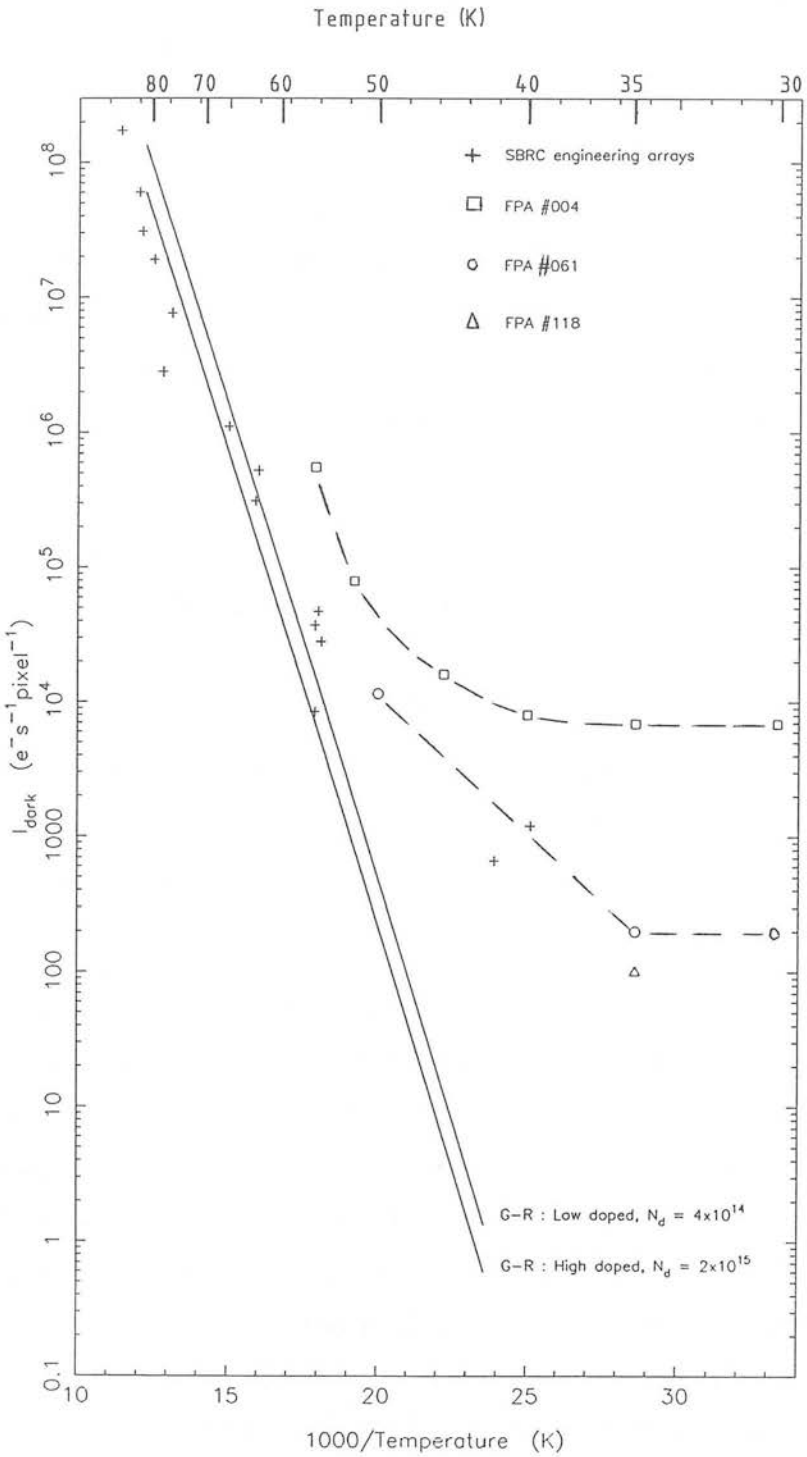


Figure 2.3 : Dark current versus temperature. Generation–recombination current is plotted, together with data from a number of high–doped SBRC InSb engineering arrays (40mV reverse bias). FPA #004 is an engineering–grade array used in IRCAM which suffered from an on–chip photon source and a decrease in quantum efficiency with array temperature. FPA #061 and FPA #118 are the first science–grade arrays used in IRCAM.

As a result of this data SBRC were only prepared to meet a specified dark current of less than 25fA at 50K ($25x10^{-15}$A or $1.5x10^5$ electrons s^{-1}) for the 62x58 InSb+DRO array. However, it was clear that cooling the array to temperatures less than 50K *would* result in significant reductions in dark current. Consequently the IRCAM design specification called for control of the detector array at any temperature in the range ~10–50K. The lower limit reflected the fact that single element InSb detectors were known to work very well at liquid helium temperatures (4K). The requirement for a *variable* temperature range reflected unknowns in the operation of the Si multiplexer at low temperatures, in particular the phenomenon of 'charge freeze–out' in the region 20–30K (Yang 1978).

A limit on the temperature stability can be estimated from the strong dependence of G–R current on n_i. By differentiating Equation 2.13 it follows that :

$$\left(\frac{dI}{I}\right)_{g-r} = \frac{dT}{T}\left(\frac{3}{2} + \frac{1400}{T}\right) \tag{2.14}$$

At 35K $dI/I \simeq dT$, and so to achieve low noise $dT \simeq \pm 0.1$K. Since at temperatures below 50K the dependence of dark current on temperature is less than the theoretical G–R dependence, a temperature stability of ± 0.1K can be regarded as a desirable lower limit.

The initial series of arrays used in IRCAM suffered from anomalously high "dark" currents. Focal plane array number 004 (FPA #004) is given as an example in Figure 2.3. (Most of the images discussed in Chapter 6 were obtained with this device.) Qualitatively the dark current curve mimics the characteristics of earlier devices, but with considerably higher currents. By placing a cold cap directly over the detector, it was established that the high signals were not due to photon leakage in the cryostat, and probably originated in the array itself. Measurements of the quantum efficiency (scaled from the system throughput) using standard stars, confirmed other reports (eg. Fowler *et al.* 1987) that the quantum efficiency rapidly decreased below ~50K before leveling off at ~40K. By not applying voltage to a column in the array multiplexer which contained several damaged MOSFETS, SBRC successfully reduced the dark current in FPA #004 to expected levels. It was concluded that the observed dark current characteristic was, in fact, due to the combined effects of the damaged MOSFETS acting as a "photon source" and to the fall of quantum efficiency with temperature.

To obtain the best performance from FPA #004 it was decided to optimise the operating temperature for best S/N in high–resolution (0.6$''$/pixel) broad band K imaging ($\lambda/\Delta\lambda$ =5 at λ =2.2 μm). For integration times long enough to make the read noise negligible :

$$\frac{S}{N} = \frac{\eta(T) I_{sky} t^{\frac{1}{2}}}{[2\eta(T) I_{sky} + 2I_{dark}(T)]^{\frac{1}{2}}} \tag{2.15}$$

where η(T) is the responsive quantum efficiency as a function of temperature T (from Figure 2.4)

$I_{dark}(T)$ is the dark current as a function of temperature (from Figure 2.3)

I_{sky} is the estimated sky background at broad band K in the high–resolution imaging mode (4x10^4 e$^-$s^{-1}pixel^{-1}, see Figure 2.6)

t is the integration time

From the plot of S/N against operating temperature (see Figure 2.5), the optimum detector temperature was found to be 49K.

Various explanations have been put forward to explain the fall of quantum efficiency with temperature (ie. the shortening of the minority carrier diffusion length in the n–type bulk when these devices are cooled below 50K). There is some theoretical evidence (Fowler *et al.* 1987 and references therein) to suggest that decreased doping can maintain the quantum efficiency. Another explanation is the absorption of carriers in impurity states contained in incorrectly processed InSb (a so called 'poisoned boule'). SBRC have now produced low–doped arrays in which the quantum efficiency is maintained at low temperature. However, due to the small statistical base it is not yet known whether this is a result of the low doping or to the arrays originating from a purer batch of InSb (Alan Hoffman,private communication 1988). The current IRCAM arrays (FPA #061 and FPA #118) are low–doped devices with very good low temperature quantum efficiencies and low dark currents (see Figure 2.3). Due to abnormalities in the functioning of the multiplexer at lower temperatures (<30K) these arrays operated at 35K. There is evidence that the performance of the multiplexer does recover below 20K, but further testing needs to be done at these temperatures.

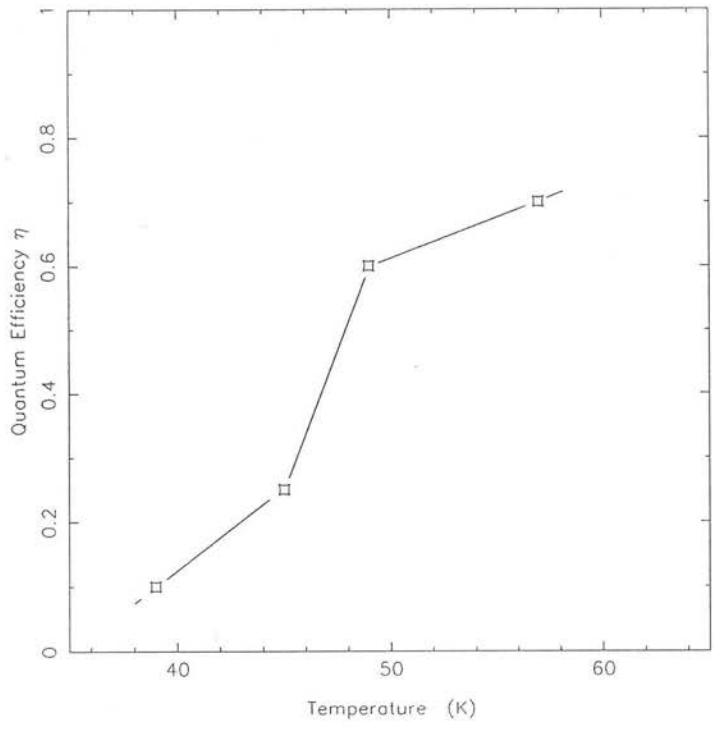


Figure 2.4 : Quantum efficiency η versus temperature for FPA #004.

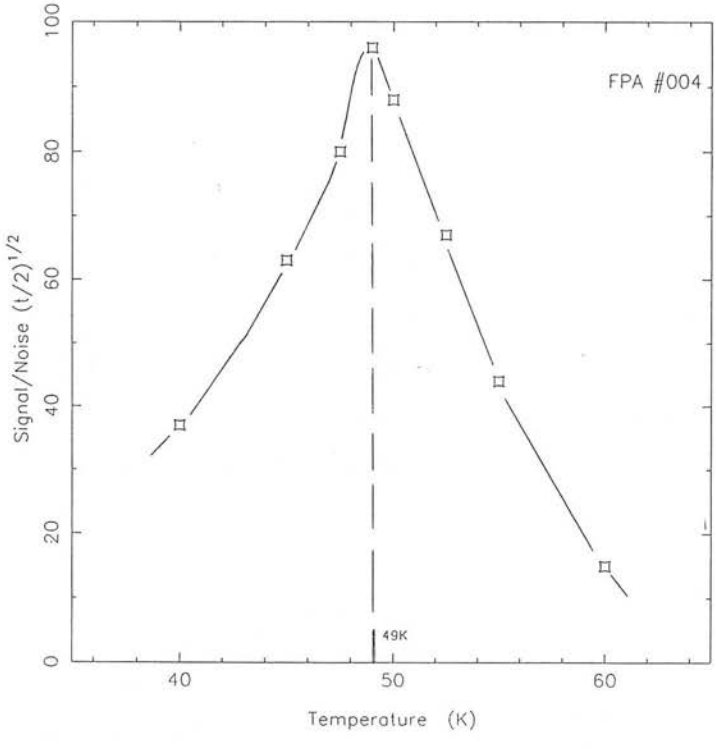


Figure 2.5 : Signal to noise versus temperature for FPA #004. Focal Plane Array 004 lost quantum efficiency when it was cooled. By optimising quantum efficiency and dark current as a function of temperature for maximum S/N in broad band (K) 2.2 μm imaging, the optimum operating temperature of the array was found to be 49K.

2.4 Background photocurrent

There are two main sources of background photocurrent :

- I_{sky}, which includes thermal and non–thermal emission from the atmosphere and telescope
- I_{instr}, which is due to thermal emission from the instrument or camera

Up to wavelengths of $\sim$2.5 μm I_{sky} is dominated by non–thermal emission, mostly airglow, from the atmosphere above the telescope, and is therefore not amenable to countermeasures. Beyond $\sim$2.5 μm thermal emission from the telescope and atmosphere dominates. Several techniques are employed to minimise thermal emission from the telescope. Where actively cooling the telescope is not practical, it is important to keep highly emissive structures out of the field of view. Highly reflective optics, as well as having low signal losses, also have correspondingly small emissivities (see Appendix A and Equation A.13). Applying Kirchoff's law it can be shown that (see also Equation 2.17) :

$$I_{sky} \propto (1 - \rho_1\rho_2....\rho_n) + \text{scattering} \tag{2.16}$$

where ρ_n is the reflectivity of the n^{th} mirror and the total emissivity $\varepsilon = 1 - \rho_1\rho_2....\rho_n$. Consequently, background can be minimised by maximising the reflectivity and reducing the number of mirrors. The scattering component arises from infrared radiation scattered into the optical path. The telescope and camera optics, and the measures taken to prevent background flux scattering into the optical path, will be discussed in Chapter 3. In the case of UKIRT + IRCAM there are five warm mirrors :

1. Primary mirror, $\rho_1 = 0.91$
2. Secondary mirror, $\rho_2 = 0.96$
3. Dichroic beam–splitter, $\rho_3 = 0.96$
4. Folding mirror, $\rho_4 = 0.98$
5. Collimating mirror, $\rho_5 = 0.98$

Resulting in a total emissivity of $\varepsilon = 0.20$. Of these only the folding and collimating mirrors could possibly be cooled, reducing the telescope emissivity to $\varepsilon = 0.16$.

Because of the additional complication required for this limited gain, cooling was not considered worthwhile. (The other warm optical component is the cryostat window. Since its small transmission loss is due to reflection, not absorption, it can be considered to have a negligibly small emissivity.) McCaughrean (1988) discusses a telescope and atmosphere model in detail and concludes that in the near–IR thermal emission from the atmosphere is small compared to thermal emission from the telescope (typically at 253K and 273K respectively).

The peak photon emission from a 273K blackbody occurs at 13.5 μm and so a 1–5 μm instrument detects the steep exponential rise characterising the Wien part of the 273K blackbody function. This means that a 2.2 μm bandpass filter such as the K filter will remove the large number of 273K photons emitted towards the detector's long wavelength limit ($\simeq$5.1 μm). Within its passband the filter will typically have a transmittance of about 70%, the corresponding 30% loss is due to absorption, resulting in an emissivity of 0.3, which is greater than the total emissivity of the telescope and the warm optics (0.20). Out of band the transmittance is very low due to reflection, resulting in a correspondingly low emissivity of perhaps 0.01. Even though the emissivity is small, significant numbers of unfiltered photons will be thermally emitted towards the detector by the filter unless it is cooled. Since the filter is essentially a flat mirror "out–of–band", it efficiently reflects photons from the surroundings, which consequently also require cooling. For the instrument background photocurrent (I_{instr}) to be negligible compared to the sky background (I_{sky}), it will be shown that the 'cold' optics and their enclosure must be cooled to <85K.

2.4.1 Sky background photocurrent

Thermal background photocurrent from a telescope at temperature T and emissivity ε is given by :

$$I_{sky} = (A_{tel}\Omega_{pfov})_{proj}\ G \int \varepsilon(\lambda)\eta(\lambda)N(\lambda,T)d\lambda \quad \mathrm{e^{-}s^{-1}pixel^{-1}} \tag{2.17}$$

where $\eta(\lambda)$ is the responsive quantum efficiency
G is the photoconductive gain – the number of electrons collected per photoelectron generated (typically 1 for photovoltaic detectors and 0.5 for photoconductors)

The Planck radiation function is given by :

$$N(\lambda,T) = \frac{2c}{\lambda^4}\frac{1}{[\exp(hc/\lambda kT)-1]} \text{ photon m}^{-2}\ \mu\text{m}^{-1} \tag{2.18}$$

where c,λ and k have their usual meanings, and the projected area solid angle product is :

$$(A\Omega)_{proj} = \int^{\phi}\int^{\theta}\int^{A} \cos\theta \sin\theta dA d\theta d\phi \ = \ 2\pi A\frac{\sin^2\theta}{2} \tag{2.19}$$

where A_{tel} is the effective collecting area of the telescope

Ω_{pfov} is the pixel field of view (pfov) on the sky in sr

ϕ=2π for a conical field of view

θ is the half angle of the cone

Note that in the optical path $A\Omega$ (also known as éntendu or throughput, see Boyd 1984) is conserved. For a relatively narrow bandwidth ($\Delta\lambda/\lambda$) the integral in Equation 2.17 may be replaced by $(\Delta\lambda/\lambda)N(\lambda,T)$. Using this equation and estimates of the 1–2.5μm airglow emission (adapted from McCaughrean 1988), Figure 2.6 shows the sky background photocurrent expected for broad band imaging (spectral resolution $R = \lambda/\Delta\lambda$=5), narrow band imaging ($\lambda/\Delta\lambda$=100) and imaging through a Fabry–Pérot étalon ($\lambda/\Delta\lambda$=3,000) on UKIRT at 0.6$''$/pixel.

2.4.2 Instrument background photocurrent

The thermal background photocurrent generated in a detector pixel by an instrument chamber at temperature T and emissivity ε is given by :

$$I_{instr} = (A_{pixel}\Omega_{pfov})_{proj}\ G\int \varepsilon(\lambda)\eta(\lambda)N(\lambda,T)d\lambda \text{ e}^{-}\text{s}^{-1}\text{pixel}^{-1} \tag{2.20}$$

Inside the instrument photons are un–filtered and the detected signal therefore depends upon $\eta(\lambda)$. Figure 2.7 is a plot of I_{instr} against chamber temperature (effectively a blackbody enclosure with ε=1), for pixels with a 2π steradian field of view. Four curves are shown corresponding to hypothetical detectors having unit quantum efficiency between the given wavelength limits. Similarly Figure 2.8 shows the I_{instr} generated in the SBRC 62x58 InSb array, assuming a 'top hat' quantum efficiency of $\eta(\lambda)$=0.7 for λ=1.0–5.0/5.2μm. The plot is shown for two cut–off wavelengths (λ_c) because of uncertainties in the precise cut–off ($\eta(\lambda_c) \rightsquigarrow 0$), and the convolution of this effect with the exponential increase in the number of detectable photons as wavelength

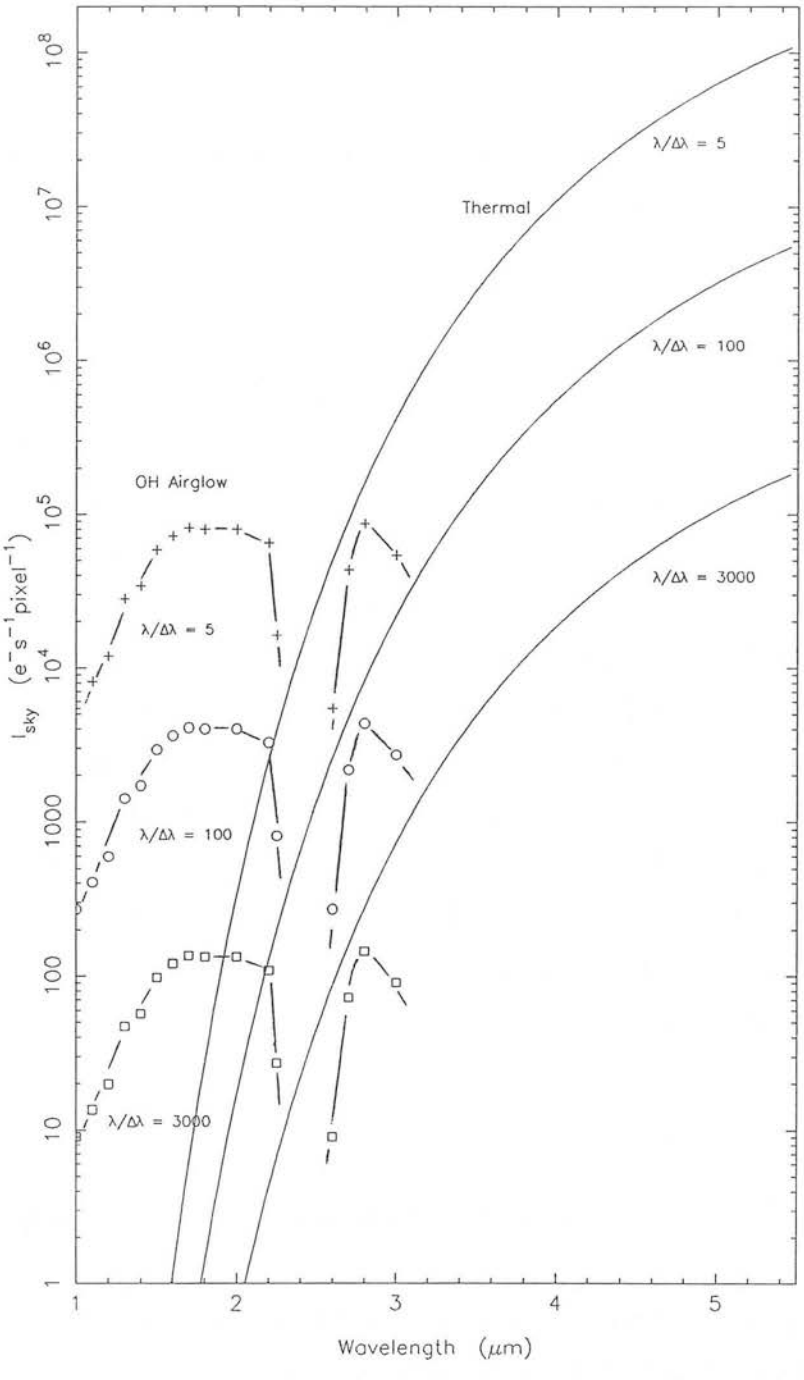


Figure 2.6 : Sky background photocurrent versus wavelength. The photocurrent generated in a 0.6″ sq. detector pixel due to emission from the sky as observed on UKIRT (collecting area 10m^2) is plotted against wavelength for broad band ($\lambda/\Delta\lambda = 5$), narrow band ($\lambda/\Delta\lambda = 100$) and Fabry–Pérot imaging ($\lambda/\Delta\lambda = 3000$). The shown thermal emission is from a telescope at a temperature of 273K and with an effective emissivity of 0.2. The OH airglow emission is adapted from McCaughrean (1988). At high spectral resolution the airglow emission is highly structured. The line shown joining the points is a guide, it is not intended to suggest a continuum. For detailed quantitative work the emission line data should be smoothed to the required spectral resolution. A detector quantum efficiency of η=0.7 has been assumed.

increases. In practice the unfiltered pixel field of view is restricted by a baffle at the same temperature as the detector, which is usually much less than that of the enclosure temperature. The reason for baffling the detector is to shield it from possible local hot–spots in the enclosure (eg. hot wires, loosened thermal insulation). Figure 2.9 is a plot of the instrument background photocurrent due to a 273K blackbody as a function of half viewing angle, from which it is clear that only a very small part of a warm object need be seen unfiltered by the detector to become a significant source of background.

The importance of cryogenically cooling instrument enclosures when observing in the IR is amply demonstrated in Figure 2.7. Although these plots must be scaled for pixel size and quantum efficiency, these are only second order effects as far as the enclosure temperatures are concerned. Because the sky background is dependent upon the bandwidth and upon the instrument application (imaging or spectroscopy), we will compare the instrument backgrounds to the device dark currents. Optical CCDs (0.3–1.1 μm,~20 μm sq. pixels) typically have dark currents of ~0.1 e^-s^{-1} and the photoelectron flux at these wavelengths from a room temperature enclosure is at most an order of magnitude less than this, so cooling is not required. A 1–2.5 μm array detector such as the Rockwell International 64x64 CMT array (50 μm sq. pixels) has a dark current of $\sim$ 10 e^-s^{-1} (Rode *et al.* 1987) and consequently requires an enclosure cooled to <165K. With a dark current of ~100 e^-s^{-1} the 1–5.1 μm SBRC 62x58 InSb array (76 μm sq. pixels) requires an enclosure cooled to <85K (see the following paragraph), while the 1–18 μm SBRC 62x58 Si:Ga array (76 μm sq. pixels), with a reported dark current of $\sim$ 1000 e^-s^{-1} (Lamb *et al.* 1987) must operate inside an enclosure cooled to <30K.

2.4.3 Implications for enclosure temperature

As already indicated, the instrument enclosure must be cooled to a temperature such that the photocurrent it generates in the detector is much smaller (say ~5 times less for a 10% degradation in the S/N) than the dark current or sky background photocurrent, whichever is greater. In such a situation the associated noise component (see Equation 2.9) is negligible. The pre–delivery (1985) detector dark current data for IRCAM (see Figure 2.3) indicated that performances of not less than ~100 e^-s^{-1} could be expected. In this case cooling the enclosure to give $I_{instr} < 20$ e^-s^{-1} would be adequate.

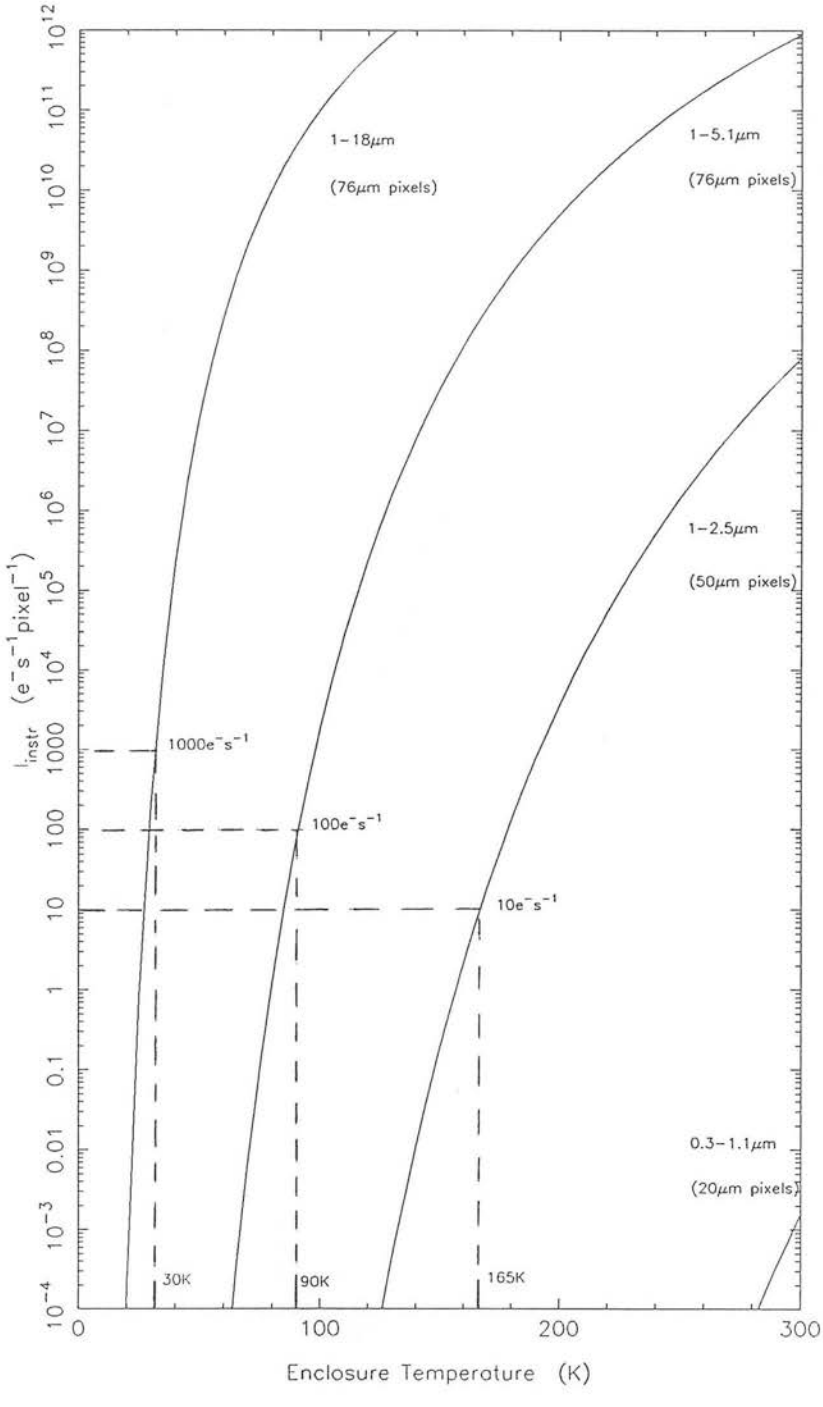


Figure 2.7 : Instrument background photocurrent versus instrument chamber temperature. The photocurrent generated in hypothetical detector pixels due to thermal emission from a blackbody enclosure is shown. The detectors are assumed to have top hat quantum efficiencies of η=1 between the given wavelength limits, 2π sr fields of view, and pixel sizes as shown. These hypothetical detectors correspond approximately to real devices (see Section 2.4.3).

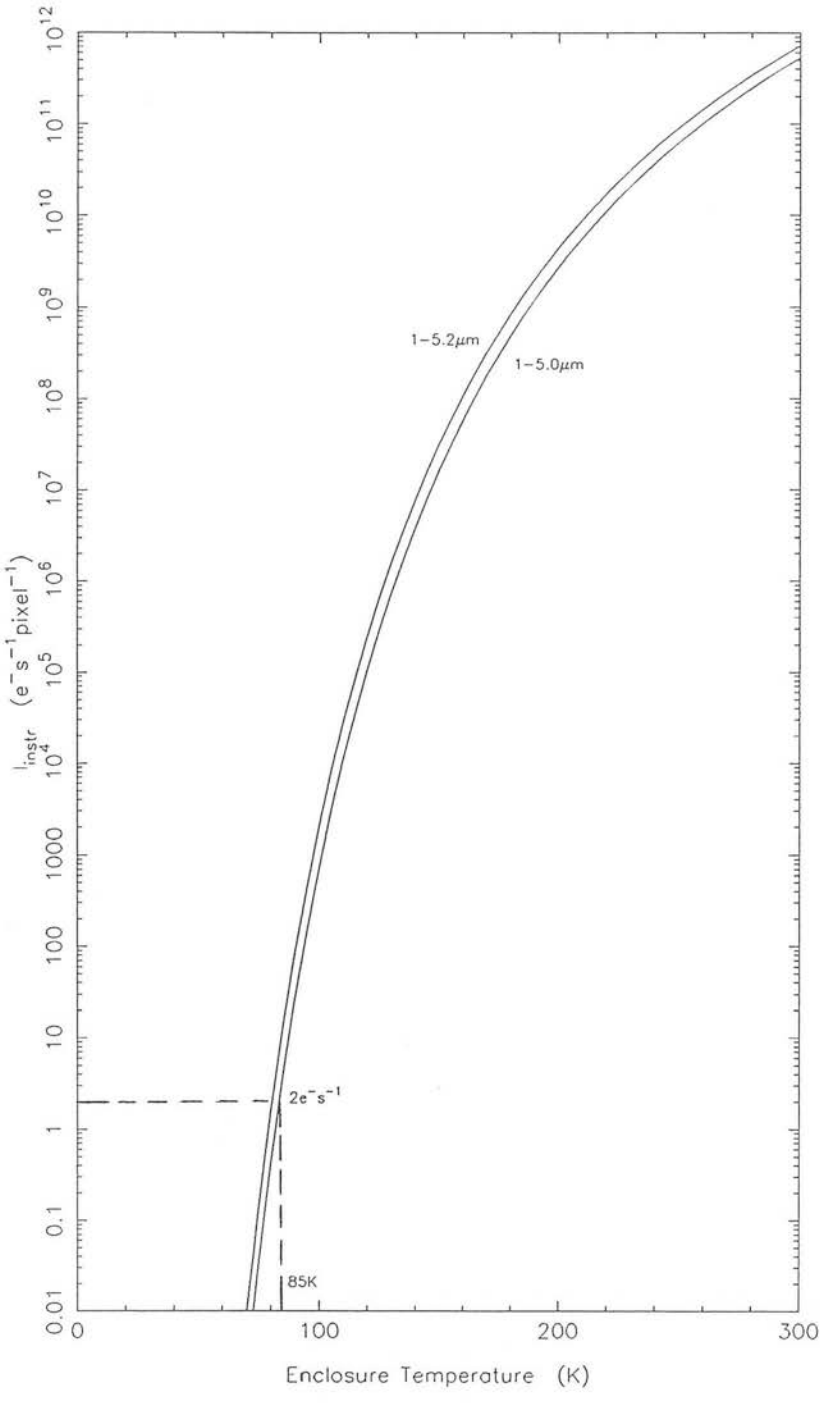


Figure 2.8 : Instrument background photocurrent generated in the SBRC 62x58 InSb array versus instrument chamber temperature. The detector is assumed to have a top hat quantum efficiency of η=0.7 between the given wavelength limits and a 2π sr field of view.

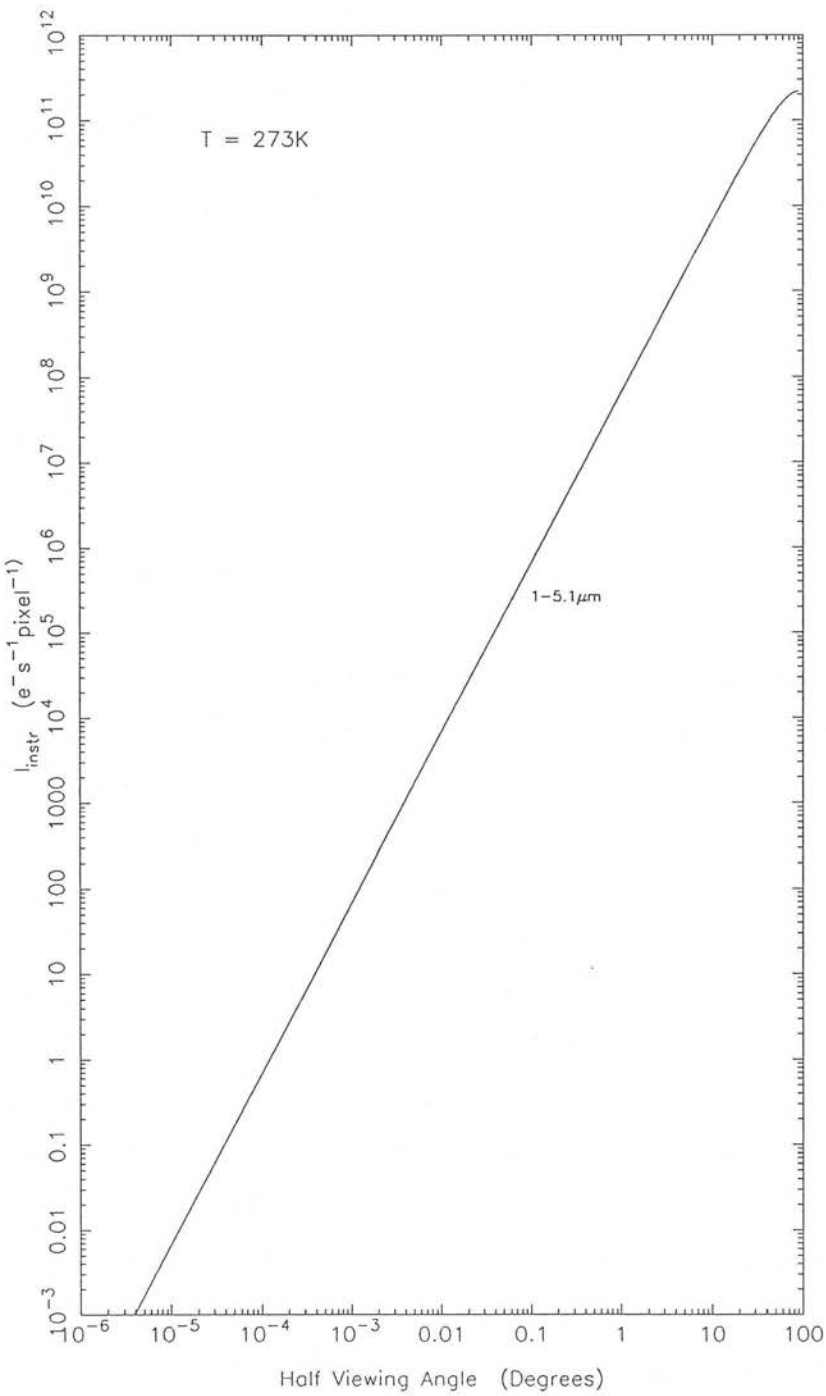


Figure 2.9 : Instrument background photocurrent due to a 273K blackbody versus half viewing angle. The detector is assumed to have a top hat quantum efficiency of η=0.7 from 1–5.1 μm.

From Figure 2.8 an enclosure temperature of <90K would achieve this. Alternatively, if the dark current can be cooled out completely, then the minimum background becomes the 10 e^-s^{-1} obtained by FP imaging ($\lambda/\Delta\lambda$=3000) at 1 μm, in which case $I_{instr} < 2\ e^-s^{-1}$, requiring an enclosure temperature of <85K. Consequently achieving an enclosure temperature of <85K was taken as the camera design specification. (The current arrays have dark currents of ~100 e^-s^{-1} which means that the camera is dark current limited for narrow band ($\lambda/\Delta\lambda$=100) imaging at <1.2 μm and FP imaging ($\lambda/\Delta\lambda$=3000) at <3 μm.)

2.5 Choice of cooling system

2.5.1 Optical constraints

The cooling system must be capable of maintaining the detector temperature and enclosure temperature while subject to inevitable heat inputs. The size of the enclosure is determined by the optical design, which in turn is driven by the scientific goals of the instrument and the usual practical compromises (risk, cost etc.). Because IRCAM is a camera, it has a large array field of view on the sky (Ω_{afov}), and since it is designed for use on a large, 4m class, telescope (A_{tel}), $A_{tel}\Omega_{afov}$ is large. So that the camera can have a high spectral and spatial resolution, it is designed to image through narrow band filters (R~100) and a Fabry–Pérot étalon (FP). To accomplish these objectives a collimator/camera scheme is adopted in which the f/36 beam emerging from the telescope focal plane is first collimated by a curved mirror, and then passed through the filters, before being re–imaged onto the array by a cold lens. In this approach the collimator also forms an image of the telescope entrance pupil (ie that part of the primary mirror "seen" by the undersized secondary mirror in the case of an optimised IR telescope) inside the camera to act as a "cold stop"; this stop is essential for rejecting off–axis background flux. The cold stop (also called a Lyot stop) is the narrowest point in the collimated beam bundle and as such is the best place for the cold filters. It is also advisable to keep it as close as possible to the cryostat window, so that the window's diameter can be minimised, which also reduces the heat input to the cryostat (although good baffling can also reduce the heat input considerably, see Chapter 3.3). Since the FP cannot be operated cold it is placed close to the window where the 'warm' beam is

narrowest.

In general, the larger the Lyot stop the bigger the optical components must be. Inside the camera the filters and window must be wider to accommodate the beam, and due to $A_{tel}\Omega_{afov}$ conservation at the cold camera lens, the wider the Lyot stop, the longer the back focal length, f_{cam} :

$$f_{cam}^2 = \frac{\pi A_{array}}{4(A_{tel}\Omega_{afov})}\,\phi_{Lyot}^2 \tag{2.21}$$

where A_{array} is the area of the detector array and ϕ_{Lyot} the diameter of the Lyot stop. Individual filters must be at least as wide as the Lyot stop, and if n filters are mounted in a wheel , the diameter of the filter wheel, ϕ_{wheel}, is :

$$\phi_{wheel} \sim \frac{n\ \phi_{Lyot}}{\pi} \tag{2.22}$$

Hence it follows that the size of the cold enclosure scales with the diameter of the Lyot stop. In effect the Lyot stop determines the radiative heat input to the cold enclosure, since this is proportional to the surface area of the enclosure.

When imaging through an FP it is desirable to minimise the number of interference fringes across the image and indeed for the central fringe to subtend the entire field of view. For this condition to be satisfied, the beam needs to be of finite width for a given spectral resolution, the wider the beam the wider the field of view. This is known as the Jacquinot criterion (see Equation 3.10), and for a given spectral resolution the diameter of the Lyot stop is proportional to the field of view. Another effect of a larger Lyot stop is that the collimator focal length is longer , which results in reduced aberrations, and relaxed external spacing constraints. In summary, the diameter of the Lyot stop is a compromise between the need to keep the camera head small and the astronomical goals of the optical design.

The IRCAM optical design (see Chapter 3.2) has resulted in a Lyot stop diameter of 19mm, 10% oversized on the actual pupil image to allow for flexure and alignment. Two filter wheels with 10 positions each offer a wide real–time choice of filters. As it is necessary to have the capability to place two filters in the optical beam at the same time (eg. for the polarisation mode in which a cold polariser plus a filter is selected), the two filter wheels are located as shown in Figure 3.6. Equations 2.19 and 2.20 applied to the high–resolution imaging mode ($39\,\widehat{''}$x $36\,\widehat{''}$ field of view), give an area in the plane of the filter wheels of 30cm x 20cm and a back focal length (f_{cam}) of 12cm, (this

is the *shortest* focal length required to sample properly a 1.2″ FWHM stellar seeing disk). When space is allowed for, amongst other things, the detector assembly and cold radiation shields, a box–shaped container of dimensions ~ 30cm x 30cm x 30cm is required to cover the cold components. The heat radiated onto an enclosure of these dimensions at a temperature of <85K from 290K is ~10W (see Chapter 4.5).

2.5.2 Cryogenic constraints

From the foregoing discussion we can now list the constraints which the cryogenic design must meet. The MOSFET array multiplexer is known to dissipate about 10mW (Fowler *et al.* 1986) and when allowance is made for the other heat inputs to the detector assembly (see Chapter 4.5), we can expect a heat load of ~50mW at the operating temperature. In summary :

- **Detector Assembly :**
  Temperature 10–50K
  Stability ± 0.1K
  Heat load ~0.05W

- **Cold Enclosure :**
  Temperature <85K
  Heat load ~10W

2.5.3 Cooling options

Cryogenic cooling can be achieved in a variety of ways (see for example Wolfe and Zissis 1978, Barron 1985) and may be classified as follows, (although some systems, particularly those requiring cooling below ~1K, are hybrids) :

1. Liquid or solid cryogens – these are open cycle expendable systems in which cooling comes from vapourisation of a liquid stored in subcritical or supercritical state, or from sublimation of a solid.

2. Open cycle Joule–Thompson (J–T) coolers – in which cooling results from allowing a gas stored at high pressure to undergo J–T expansion through a valve.

3. Passive radiators – which cool systems to cryogenic temperatures by radiation to a low temperature, usually deep space.

4. Thermoelectric coolers – which use the Peltier effect. When two dissimilar metals are connected in series with a voltage source, one junction will be cooled while the other will be heated.

5. Closed cycle refrigerators – these include mechanical refrigerators based on a variety of thermodynamic cycles (eg. the Gifford–McMahon refrigerator, see Walker 1983), and more exotic magnetic and dilution refrigerators in which expansion of the working fluid of a mechanical refrigerator is replaced by demagnetisation and phase separation respectively (necessary to reach temperatures $\ll$1K).

Of these cooling techniques only liquid or solid cryogen systems, and mechanical refrigerators have sufficient cooling power at the required temperatures.

2.5.4 Mechanical refrigerators

The Model 350 C CRYODYNE$^{(R)}$ Cryocooler made by CTI–CRYOTRONICS is an example of the type of mechanical refrigerator considered for cooling IRCAM. This particular cryocooler operates on the Gifford–McMahon thermodynamic cycle in which the compressor and the expander assembly (containing two cold stages) are separated from one another, and are interconnected by two helium gas lines. The expander requires a small motor to move a gas displacer up and down, and to operate the valves in the expander head. The presence of valves in both the compressor and expander isolates them from one another and allows the interconnecting lines to be of almost any length since their volumes are effectively isolated from the thermodynamically active volumes in the system. Separation of the compressor and expander removes the principal source of vibration from the region of cooling, however smaller vibrations due to the displacer still remain. Typical refrigeration capacities of the two stage 350 C Cryocooler are shown in Figure 2.10 (see CTI–CRYOTRONICS 1984), and are more than capable of providing the required cooling for about one year without shutdown for maintenance. Such performances result in considerable cost savings over expendable,

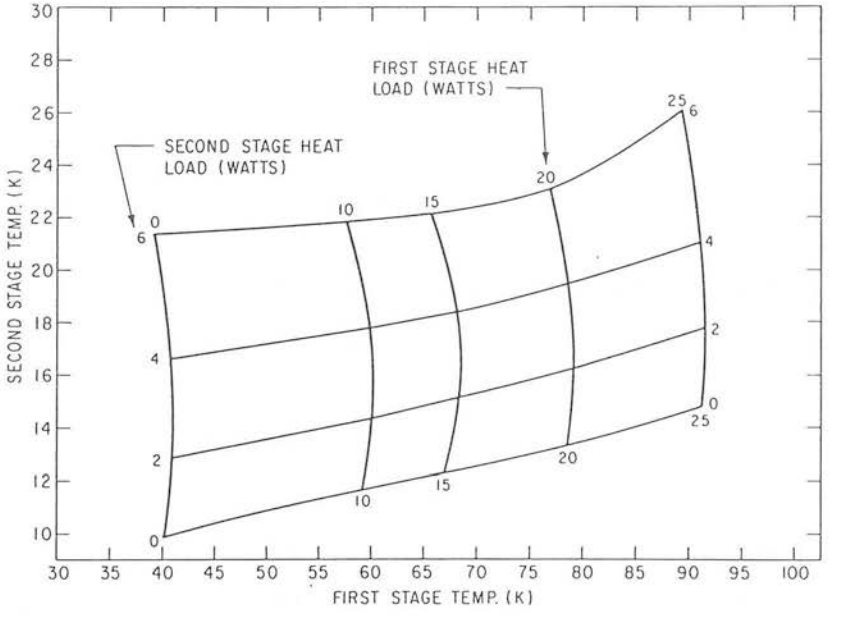


Figure 2.10 : Typical cooling capacity of the Model 350C Cryodyne^{R} Cryocooler (60H$_{Z}$).

open cycle liquid and solid cryogen systems and so they were seriously considered for IRCAM. The temperatures of the cold stages are heat load dependent, which means that temperatures in the range 10–50K can be obtained by feedback control of a heater mounted onto the second stage. Cooling is independent of orientation.

As already mentioned, one of the inherent problems with mechanical refrigerators is vibration. UKIRT is operated with a cluster of up to 4 instruments at the base of the telescope, all of which must be operational. Even if vibration proved unimportant for the camera, it could still pose serious problems for the sensitive bolometer systems, where small vibrations can cause microphonic noise. Microphonics arise from current and capacitance changes when electrical wires and components move in electric fields. In principle vibration–isolation mountings can be used to reduce these problems. Another concern is the cooling time which can become very long when the cold stages are connected to large masses such as the IRCAM camera head.

In 1984 when the decision had to be made on which cooling system to use , ROE had no experience in the use of mechanical refrigerators. Consequently, it was decided to adopt the well–understood liquid cryogen dewars while independently trying to gain experience with mechanical refrigerators . At the time of writing (March 1988), the first trials of a 350 C Cryocooler mounted, via a vibration–isolation mounting, as one of the instruments in the cassegrain cluster on UKIRT, have been successfully completed. It will still be some time before such a cooler is used in a common–user instrument, however.

2.5.5 Liquid/solid cryogen dewars

The terms "dewar" and "cryostat" are sometimes used interchangeably, but in this thesis a dewar will be understood to mean "a container for storing cryogenic fluids", and a cryostat "a cryogenic instrument consisting of a dewar and cooled instrumentation".

Cryogenic fluids can be stored in equilibrium with their vapours at subcritical pressures (two phases), or at supercritical pressures and higher temperatures, as single phase homogenous fluids. Single phase storage finds direct application in zero–G environments where gravity is not able to keep a cryogenic liquid in contact with a cold sink, to which cooled instrumentation is thermally connected for example. There are

Table 2.1: **Selected properties of cryogenic fluids**

Saturated – liquid *Property*	*Liquid* *Helium 4*	*Liquid* *Hydrogen*	*Liquid* *Neon*	*Liquid* *Nitrogen*	*Liquid* *Argon*
Normal boiling point temp. (K)	4.2	20.3	27.1	77.4	87.3
Critical temperature (K)	5.2	33.2	44.4	126.1	150.7
Critical pressure (atm)	2.3	13.0	26.2	33.5	48.3
Triple – point temp. (K)	–	13.9	24.5	63.2	83.8
Triple – point press. (atm)	–	0.07	0.43	0.13	0.68
Density ($kg\ m^{-3}$)	125	70.8	1206	807	1394
Latent heat ($kJ\ kg^{-1}$)	20.9	443	85.9	199.3	161.9

two basic types of open cycle two phase cryostats based on dewars. In the first and simplest, instrumentation is built into a dewar and cooling is by direct thermal contact with the storage vessel, this is the familiar bath cryostat. In the second, a liquid–feed cryostat, cryogenic fluid is fed to the instrumentation from a remote dewar through transfer lines. The performance available from cryostats depends largely on the properties of the cryogenic fluids. Table 2.1 summarises the relevant properties of cryogenic fluids covering the temperature range of interest. By varying the vapour pressure the temperature of the cryogen can be changed. The lower the pressure, the lower the temperature; the limiting temperature depends on the ability of the pumping system to remove vapour, and the parasitic heat leak into the cryogen generating the vapour. It is standard procedure to reach temperatures below 77K by attaching an evacuation pump to a liquid nitrogen (LN_2, 1 atm boiling point 77.4K) fill/vent port. In this manner the vapour pressure is easily reduced to below the triple point pressure (0.127 atm, T_3=63.2K) and solid nitrogen (SN_2) temperatures of 47K are typically attained.

To avoid excessive loss of cryogens it is essential to minimise heat loads. In general cryostats utilise vacuum, multilayered insulation (MLI) blankets, cryogen– and vapour–cooled shields, together with rigid support structures of low thermal conductivity to achieve the low heat loads necessary for long hold–times. Although in principle a dewar provides a fixed temperature sink, in practice cooled components are connected to the

temperature sinks by structures with finite thermal resistances and so any heat load will cause a finite temperature difference.

2.6 The IRCAM cryostat

The IRCAM cryogenic container is a dual vessel (two temperature stage) LN_2/LHe cryostat. In principle the cryogenic performance can be met by a single cryogenic temperature stage which can be varied from $\sim 10 - 50$K. This temperature range can be achieved by connecting the stage to a liquid helium (LHe, 1 atm boiling point 4.2K) vessel via a variable thermal resistance (see below). A slightly smaller range can be obtained by using liquid neon (LNe, 1 atm boiling point 27.1K). By pumping over liquid neon it is possible to reach a solid neon (SNe) temperature of 17K (0.004 atm for a 100mW heat load), and so by controlling the back pressure up to the critical pressure of 26.2 atm (tank integrity permitting), a useful temperature range of 17–44K results. Alternatively temperatures above 27K can also be attained by connecting a variable thermal resistance to a LNe vessel. A single stage cryostat must be able to sink the anticipated 10W of heat load. It is useful to express a cryogens cooling ability in terms of its density (ρ) and latent heat of vapourisation (L_V) product. In the appropriate units :

$$(\rho L_V)_{LHe} = 0.74\ W\ hr\ l^{-1}$$

$$(\rho L_V)_{LNe} = 28.8\ W\ hr\ l^{-1}$$

$$(\rho L_V)_{LN_2} = 44.7\ W\ hr\ l^{-1}$$

This means that a 10W heat load will boil away one litre of LHe, LNe or LN_2 in 0.07 hr, 2.9 hr and 4.5 hr respectively. Since LHe and LNe are considerably more expensive than LN_2, a single vessel cryostat means unrealistic operating costs and in the case of LHe an unrealistic vessel size to achieve reasonable hold–times (>330l for 1 day). It is also desirable to reach any detector temperature in the operating range within about one hour once the camera has cooled, so the low mass detector must be thermally de–coupled from the rest of the cold components and operated as a separate temperature stage. For these reasons a *dual vessel* cryostat design was chosen.

The first stage temperature in the dual vessel cryostat is fixed at 80–85K by cooling with LN_2. As well as providing the required cooling for the optics enclosure, the LN_2

vessel also cools a radiation shield which shrouds the lower temperature cryogen vessel from ambient temperature radiation. Since SBRC were not under contract to test arrays below 50K, it was necessary to explore detector performance down to ~10K. Futhermore, since LHe was already available at UKIRT to cool bolometer instruments, it was practical to use LHe as the low temperature coolant and to establish a variable temperature detector cold stage by connecting the detector assembly to the LHe vessel by a variable thermal conductance. Variable temperatures are usually produced by using a heater to increase the heat flow through a fixed thermal resistance and thereby cause a rise in temperature. This of course has the disadvantage of increasing the heat leak into the cold sink. By varying the thermal resistance/conductance instead, this is avoided. A Variable Conductance Gas Switch (VCGS) was developed specifically for this purpose and is fully discussed in Chapter 5.

The configuration of the IRCAM cryostat is shown in Figure 2.11. It consists of two main sections :

1. **The Dewar** – This is the cooling section and contains :
   - LHe vessel (4.0 litres)
   - LN_2 vessel (4.5 litres)
   - Cold radiation shields
   - VCGS
2. **The Camera Head** – This is the instrument section and contains :
   - Detector assembly
   - Focus assembly
   - Lens assembly
   - Filter wheel assembly
   - Cold baffle
   - Pre–amplifier
   - Cold optical table and radiation shields

All the cooled components are mounted on the cold optical table which breaks with tradition by being mounted from the *warm* outer camera body by stiff fibre–glass

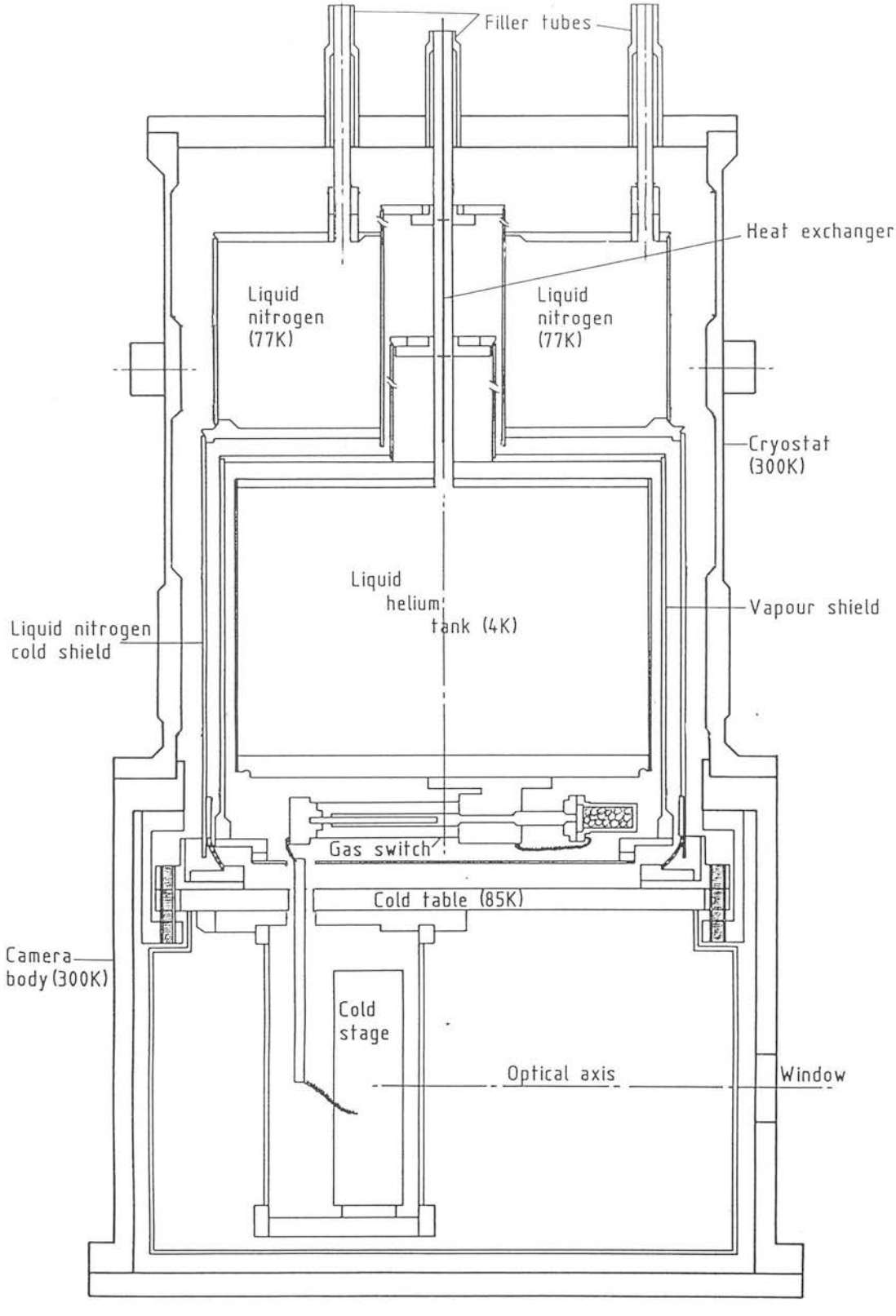


Figure 2.11 : The IRCAM Cryostat. This consists of an Oxford Instruments MD1810 dual vessel LN_2/LHe dewar connected to a customized camera head which contains the detector (mounted on the cold stage) and the cold optics (not shown).

thermal–isolation pillars. Excellent rigidity is achieved in this way (see Chapter 4.4) The cold table is cooled by a ring of flexible copper braids connected to the LN_2 radiation shield. The detector assembly is rigidly supported from the cold table by two fibre–glass pillars, which also thermally isolate it from the ~85K cold table, allowing it to operate in the range 10–50K. The latter is accomplished by thermally connecting the detector assembly (or cold stage) to the LHe cold face through a flexible copper cold finger and the VCGS. The camera head and its detailed cryogenic design are the subjects of Chapters 3 and 4.

By keeping the Dewar and Camera Head separate, this design enables the dewar to be replaced by a mechanical refrigerator at a later date with only minor modifications to the dewar/camera head thermal and mechanical interface.

Chapter 3

The Camera Head

3.1 Introduction

In this chapter we will show how the optical design of IRCAM has determined the layout of the camera head. The optical design has been driven primarily by the scientific goals required of a 1–5 μm camera, and constrained by the nature of the telescope and detector array. For example, UKIRT is configured to operate with an f/36 optical beam and has a 200 $\widehat{''}$ field of view on the sky. For accurate stellar photometry, stellar seeing disks typically 1 $\widehat{''}$ FWHM must subtend at least two pixels (Nyquist sampling). To cover all the available field of view this requires a 400 x 400 pixel array. With the 62 x 58 pixel array currently used in IRCAM, this can only be achieved by constructing a mosaic. At the other extreme we could aim for maximum field of view without the need for constructing a mosaic, at the expense of accurate stellar photometry, but with increased sensitivity to diffuse low surface brightness structures; then from $A\Omega$ conservation :

$$\theta_{array} = \frac{\phi_{tel}}{\phi_{array}} \theta_{tel} \tag{3.1}$$

where θ_{array} is the acceptance cone angle of the beam converging on to the array
ϕ_{array} is the effective diameter of the array (62x76 μm; pixels are 76 μm square)
θ_{tel} is the acceptance cone angle corresponding to a 200 $\widehat{''}$ telescope field of view
ϕ_{tel} is the effective diameter of the telescope's primary mirror (3.70m)

From this it follows that θ_{array} = 0.76rad and by inverting this we get the f–number (single symbol f/#=focal length/diameter) and the focal ratio of the last optical element is f/1.3. A general property of cameras is that the wider the field of view, the 'faster' the optics (small f–numbers, large cone angles). Optical aberrations are inversely proportional to some power of the f–number and become significant for optical components operating at ~f/1. Multiple component optical systems can be used to minimise the aberrations by dividing the amount of beam bending between its components. However, in astronomy, serious light loss cannot be tolerated, therefore the number of optical components must be minimised. Consequently the speed of the optics is a compromise between considerations of field of view and aberrations.

Ideally astronomical images should be acquired over a large range of spectral resolutions. Standard broad band filters (R~5) are matched to windows in the transmission of the atmosphere (JHKL'M) for maximum S/N in continuum emission from sources. Individual filters can be matched to particular spectral line profiles (R~100) to minimise the continuum contribution for maximum S/N, but they require matching complementary continuum filters to allow continuum subtraction. Circular variable filters (CVFs) are often used in aperture photometry but their resolutions (R~100) are determined by the size of the optical beam on the filter (typically ~1mm), which is tuned to achieve continuum subtraction from the line. Imaging spectroscopy (R~1000–10,000) is most easily achieved using a Fabry–Pérot étalon (FP) with order sorting either by a CVF or narrow band filter. At low spectral resolutions 1–2m class telescopes are competitive with 4m class telescopes, since although it takes longer to get background limited, observing time is generally more easily come by. At high spectral resolutions (R>1000) it is very difficult for the smaller ground–based telescopes to get background limited whereas the 0.5–1m cooled space–based telescopes lack the collecting area needed for good S/N at high spectral resolution. Because of this significant performance advantage, the provision of an imaging spectroscopy capability for UKIRT has been a major consideration in the design.

It is the requirement of wide field high spectral resolution imaging which imposes the most stringent constraints on the optical design. To be used effectively the FP must be placed in a parallel (or at least very slow) beam, and because of the large beam diameter (17mm, see Section 3.2.2) required to maintain a wide field of view when using the FP as a very narrow band filter, it was not practicable to incorporate a CVF

into the design. As explained in Section 3.2.1 a Collimator/Camera design provides IRCAM with a choice of 3 image scales, a selection of broad band and narrow band filters, and the option of placing a warm FP in the collimated beam. The resulting layout of the camera head is discussed in Section 3.4.

The optical design employs a combination of mirrors and lenses to transfer the flux collected by the telescope primary mirror onto the detector array. Because of the highly emissive nature of warm objects surrounding the optical beam, it is vital to prevent this thermal background flux from being scattered into the beam. In Section 3.3 we will examine the magnitude of the background flux which can be scattered into the beam, and show how a hybrid system of warm and cold imaging baffles, are implemented.

3.2 Optical Design

3.2.1 Telescope optics

Optical interfaces of the camera to the telescope occurs at the telescope focal plane and naturally requires a knowledge of the location of the focal plane, the image scale, and the diameter of the field of view on the sky. UKIRT has a 3.80m diameter f/2.5 parabolic primary mirror and is currently configured to work with a 0.315m diameter f/2.75 hyperbolic secondary mirror, which results in an f/36.15 beam converging to the telescope focal plane. The secondary mirror is small enough to allow 'sky chopping', necessary for sky background subtraction in non–integrating, a. c. coupled detectors. Since UKIRT is optimised for the IR, the secondary mirror is positioned in the converging beam from the primary mirror such that it 'sees' an effective diameter of 3.70m, ensuring that thermal flux from the highly emissive structures surrounding the primary is not transferred to the focal plane. The secondary mirror therefore defines the telescope's entrance aperture. (In optical (non–thermal) telescopes, the secondary is oversized so that no flux is lost.)

The secondary mirror forms an image of the primary mirror focus 11.388m from the secondary mirror pole; this is the telescope focal plane. A dichroic beam–splitter placed in the f/36.15 converging beam directly below the central hole in the primary mirror, splits the beam into an axial optical component and a radial IR component.

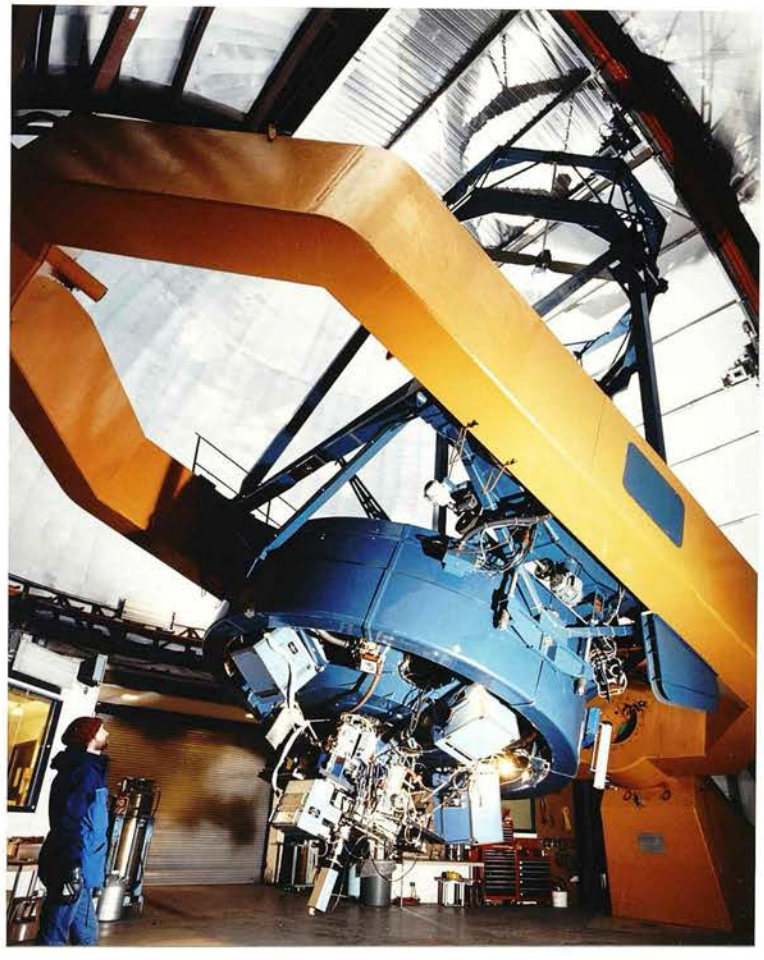


Figure 3.1 : UKIRT. The 3.8m telescope is usually configured to operate with a cluster of 4 radially mounted side–looking instruments. In this view IRCAM is shown mounted to the north port of the Instrument Support Unit (ISU2). The cryostat is seated on the Instrument Mounting Platform (IMP). Also visible are three cabinets containing electronics for the camera. The array drive electronics and A/D converter are mounted below the IMP. The Array Control Processor (ACP), sequencer and computer interface electronics are contained in a cabinet attached to the mirror cell (left). A third cabinet, also attached to the mirror cell (right), contains the drive motor control unit, together with power supplies and control electronics for the temperature controller.

By rotating the dichroic mirror the radial beam can be steered towards any one of the 4 radially mounted side–looking instruments without affecting the axial optical beam, which is used by the TV guide camera (see Figure 3.1). By differentiating the simple lens formula it follows that :

$$\text{Longitudinal magnification, } \frac{dv}{du} = \frac{v^2}{u^2} \tag{3.2}$$

where v is the distance of the secondary mirror from the focal plane (11.388m)
u is the distance of the primary mirror (virtual) focus from the secondary mirror pole ($9.503 - 8.694 = 0.809$m, see Figure 3.2)

The longitudinal magnification at the focal plane is ~200 which means that the position of the focal plane is highly sensitive to small changes in the position of the secondary mirror (a case being solar heating of the telescope top–end). A 100mm diameter inclined plane mirror is placed over the secondary mirror pole so that the highly emissive 1.00m diameter cassegrain hole in the primary mirror is not seen. Instead the plane mirror is inclined to view the sky in the primary mirror. In the appropriate units, the telescope focal plane image scale is given by :

$$\frac{1}{f\#\ \phi_{tel}} = 1.55\,\widehat{''}\text{mm}^{-1} \tag{3.3}$$

where $f/\#$ is the effective f–number of the telescope (36.15)
ϕ_{tel} is the effective diameter of the primary mirror (3.70m)

By drawing rays through the secondary mirror and the image of the primary mirror in the secondary mirror, and projecting these rays into the focal plane (see Figure 3.2), the unvignetted UKIRT f/36.15 focal plane diameter is found to be 130mm, which from Equation 3.3 is equivalent to a $200\,\widehat{''}$ field of view on the sky. This is the maximum field of view available to the camera. For a point in the centre of the field of view there is a 50mm unseen annulus inside the edge of the primary mirror which is available for chopping without going off the edge. This is equivalent to $100\,\widehat{''}$ (65mm) in the focal plane. The camera has a maximum field of $135\,\widehat{''}$ (see Section 3.2.2), which allows a maximum chop distance of only $\pm 32.5\,\widehat{''}$ before the edge of the field starts to see the highly emissive objects surrounding the primary mirror. This is a relatively small maximum chop distance for wide field instruments and so it is planned to replace the current secondary with a smaller diameter mirror to increase the available chop throw.

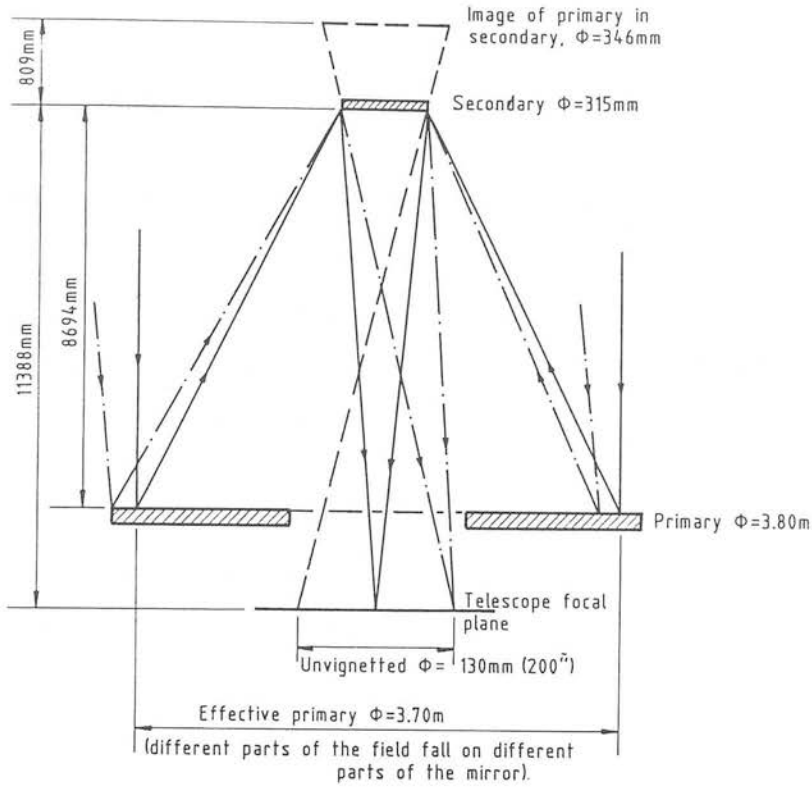


Figure 3.2 : Telescope optics. The solid lines represent rays from an object in the centre of the field of view, the dotted lines represent rays from an object at the edge of the field of view. The telescope dimensions (not shown to scale) were provided by G. Adam (private communication, 1988).

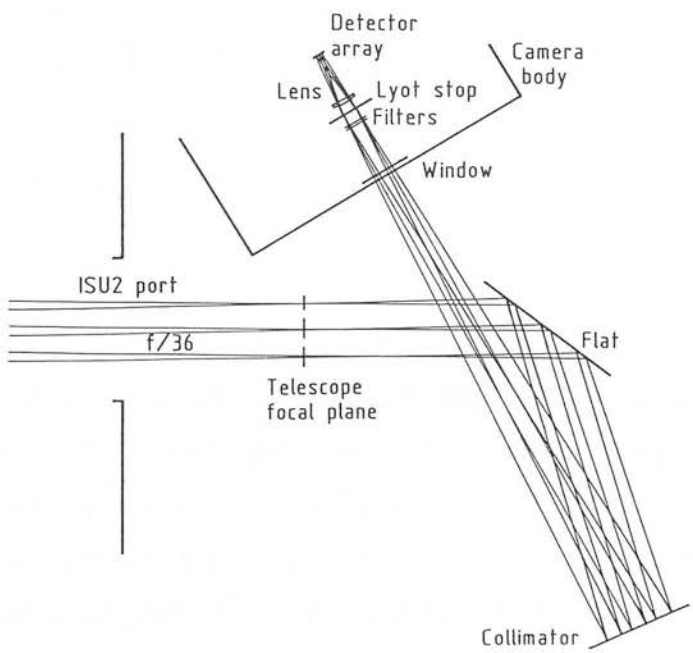


Figure 3.3 : Camera optics. The f/36 beam emerging from the telescope focal plane is folded by a plane (flat) mirror before being collimated and redirected into the IRCAM camera head, where a cold lens focuses it onto the detector array (to scale).

3.2.2 Camera optics

The detailed optical design was done by IRCAM's Project Scientist Dr Ian McLean with assistance from the author during testing and evaluation. Many optical configurations were considered and the final design was optimised by extensive ray tracing. Only those issues affecting the general layout of the camera head, and therefore its cryo–mechanical performance will be discussed here. The requirement of the optical design is to :

- Match the f/36.15 image scale ($1.53\,\widehat{''}\mathrm{mm}^{-1}$) of UKIRT
  to the 62x58 pixel (4.71mm x 4.41mm) InSb detector array to achieve :
  1. Nyquist sampled stellar photometry
     (high–resolution mode $\sim 0.5\,\widehat{''}$ pixel^{-1}, FOV $\sim 0.5\,\widehat{'}$)
  2. Wide fields of view on the sky
     (low–resolution mode $\sim 2.5\,\widehat{''}$ pixel^{-1}, FOV $\sim 2.5\,\widehat{'}$)
- Accommodate a wide wavelength range
- Provide a parallel input beam to the cryostat to give the camera a Fabry–Pérot imaging spectroscopy mode
- Form an image of the telescope entrance aperture inside the cryostat to act as a cold stop (Lyot stop)
- Minimise the number of optical surfaces to prevent serious light–loss and image ghosts

The scheme adopted to satisfy these conditions is a variant of a classical design often employed in astronomical instruments, the Collimator/Camera (see Figure 3.3). In this approach the f/36.15 beam emerging from the telescope is first collimated by an off–axis spherical mirror to produce bundles of parallel beams (of diameter equal to the image of the entrance aperture in the collimator), corresponding to different points in the telescope focal plane. The collimating mirror is placed one focal length (f_{coll}) from the focal plane and has the property that it produces a demagnified image of the telescope entrance aperture at a distance slightly greater than f_{coll} from its surface. This image is known as the Entrance Pupil, and since all the collimated beam bundles

(from any point in the field of view) must path through it, it is the narrowest part of the collimated beam. Precisely because of this, the entrance pupil is the best location for a cold stop, which baffles the IR system from off–axis thermal flux. A stop so located is called a Lyot stop. The diameter of the Lyot stop (ϕ_{Lyot}) is nominally the same as the entrance pupil and which is given by :

$$\frac{\phi_{Lyot}}{f_{coll}} = \frac{1}{f/\#} \tag{3.4}$$

where the effective f–number of the telescope $f/\#$ is 36.15

The Lyot stop is slightly larger than the entrance pupil to prevent change in signal due to flexure. To re–image the focal plane, a camera lens or mirror is placed in the collimated beam just after the Lyot stop. Since the detector is physically small, the focal plane image of the sky must be demagnified by the camera lens. The reduction in field size is given by :

$$X_{array} = \frac{f_{cam}}{f_{coll}} X_{tfp} \quad \text{where } f_{cam} < f_{coll} \tag{3.5}$$

where X_{array} is the linear dimension of the array (4.7mm)
X_{tfp} is the linear dimension of the field in the telescope focal plane
f_{cam} is the focal length of the camera lens
f_{coll} is the focal length of the collimator mirror

For example, spatial resolutions of 0.6, 1.2 and 2.4 $\widehat{''}$ pixel^{-1} are produced by demagnifications (X_{tfp}/X_{array}) of 5, 10 and 20 respectively. From Equations 3.4 and 3.5, the speed of the camera lens is given by :

$$\frac{f_{cam}}{\phi_{Lyot}} = f/\# \ \frac{X_{array}}{X_{tfp}} \tag{3.6}$$

where the diameter of the camera lens needs to be only slightly wider than the entrance pupil, and therefore the Lyot stop, to accept the beam. Consequently, the speed of the lens depends only on the desired field of view, and *not* on the individual values of ϕ_{Lyot}, f_{cam} and f_{coll}; this is just another statement of $A\Omega$ conservation. For small aberrations $f_{cam}/\phi_{Lyot} > 1$. As mentioned previously (Chapter 2.5.1), there are advantages in keeping the Lyot stop small since the optical components located near it (the cryostat window, cold filters and camera lens) may then also be kept small. Also, the focal length of the collimator is proportional to the diameter of the Lyot stop (Equation 3.4), and so the Lyot stop constrains the spacing of the collimator and vice versa.

A fundamental constraint on the diameter of the collimated beam bundle, and therefore the diameter of the Lyot stop, comes from the requirement of imaging spectroscopy using a Fabry–Pérot étalon. An FP etalon is in effect a tuneable interference filter and it can be shown that for wavelength maxima in the interference film or gap :

$$\lambda = \frac{2t}{m}\mu \cos\left(\frac{\theta}{\mu}\right) \tag{3.7}$$

where t is the thickness of the gap

μ is the refractive index of the film or gap

m is the interference order

θ is the acceptance cone angle of the incident beam in free space

If it is assumed that the reflecting boundaries are perfectly parallel and that the half–width $\Delta\lambda$ is determined only by the reflection coefficients, then the acceptance cone diameter θ which produces a wavelength shift across the central interference fringe of $\Delta\lambda$, is given by :

$$\theta^2 = \frac{8\mu^2}{R} \quad \text{where } \theta < 20° \tag{3.8}$$

or in terms of solid angle by :

$$\Omega = \frac{2\pi\mu^2}{R} \tag{3.9}$$

In essence these equations say that for the spectral resolution of an interference filter to be determined mainly by the flatness of the boundaries and their reflectances and *not* by the illumination angle of the incident beam then :

$$\Omega < \frac{2\pi\mu^2}{R} \tag{3.10}$$

This relationship has become known as Jacquinot's criterion (Meaburn 1976). When an FP (or interference filter) is placed in a collimated beam, then for *each point* in the focal plane, the acceptance cone angle is zero and there is no geometrical degradation of the spectral resolution. In the case of a narrow band filter of R$\sim$100 the cone angle can approach 20° (speed f/2.3) before the resolution is degraded. However even when the beam is collimated, collimated beam bundles from *different points* in the focal plane constitute a field of view acceptance cone angle. Consequently for the wavelength shift across an image to be less than $\Delta\lambda$ the acceptance cone angle of the collimated beam bundles must satisfy Equation 3.10. When this condition is not satisfied the image consists of a central fringe and a number of concentric rings. Because of the high

spectral resolutions, FPs are particularly sensitive to this effect, and since it is highly desirable for the entire field of view to remain within the central fringe, Equation 3.10 must be satisfied. The $A\Omega$ product of the beam through the FP is the same as that at the Lyot stop, and so for the FP to be matched to a given field of view :

$$A_{Lyot}\Omega_{Lyot} = A_{tel}\Omega_{tel} \tag{3.11}$$

and by substituting for Ω_{Lyot} from Equation 3.11 in Equation 3.10 :

$$A_{Lyot} > \frac{A_{tel}\Omega_{tel}}{2\pi\ \mu^2}\ R \tag{3.12}$$

This important result means that the diameter of the Lyot stop scales with the field of view and spectral resolution. For IRCAM the resolution is limited to R$\sim$3000 by the R$\sim$100 narrow band filters which are used as order sorters.

Optimising the various parameters has resulted in a beam diameter at the Lyot stop of 16.7mm (the stop itself is $\sim$10% oversized to allow for flexure and alignment), a collimator focal length of 610mm, and a choice of 3 spatial resolution scales, 0.6, 1.2 and 2.4 $\widehat{''}$ pixel^{-1}. The beam sizes required to satisfy the Jacquinot criterion are 7.5mm, 13.0 and 26.0mm respectively, and so only in the low–resolution 2.4 $\widehat{''}$ pixel^{-1} is the beam too small and we can expect to see interference rings. The low–resolution image scale is limited by the need to restrict the speed of the fastest lens to $f_{cam}/1.75$ because of aberrations. In the low spatial resolution mode, the unvignetted field of view is limited to a diameter of about 135 $\widehat{''}$ (56 pixels) by the 120mm diameter collimator mirror ($f_{coll}/5.1$).

Using an all mirror design has the advantage of being achromatic and the high reflectances automatically ensure a low emissivity and therefore a low thermal background if the mirrors are warm. However mirrors must be used either off–axis to avoid self obscuration, or centrally obscured. In the first case incorporating several off–axis mirror options of different focal lengths inside the cryostat to permit a choice of final image scale, has proved very difficult, while in the second case, the loss of flux is similar to transmission losses in the simpler lens configuration. The final design employs off–axis warm mirrors in the collimator and cold on–axis lenses in the camera head. To keep the design simple, the lenses are not achromatic and therefore focus is wavelength dependent. The calculated spot diagrams show that the blur circles introduced by coma, spherical aberration and astigmatism are less than one pixel at the edge of the

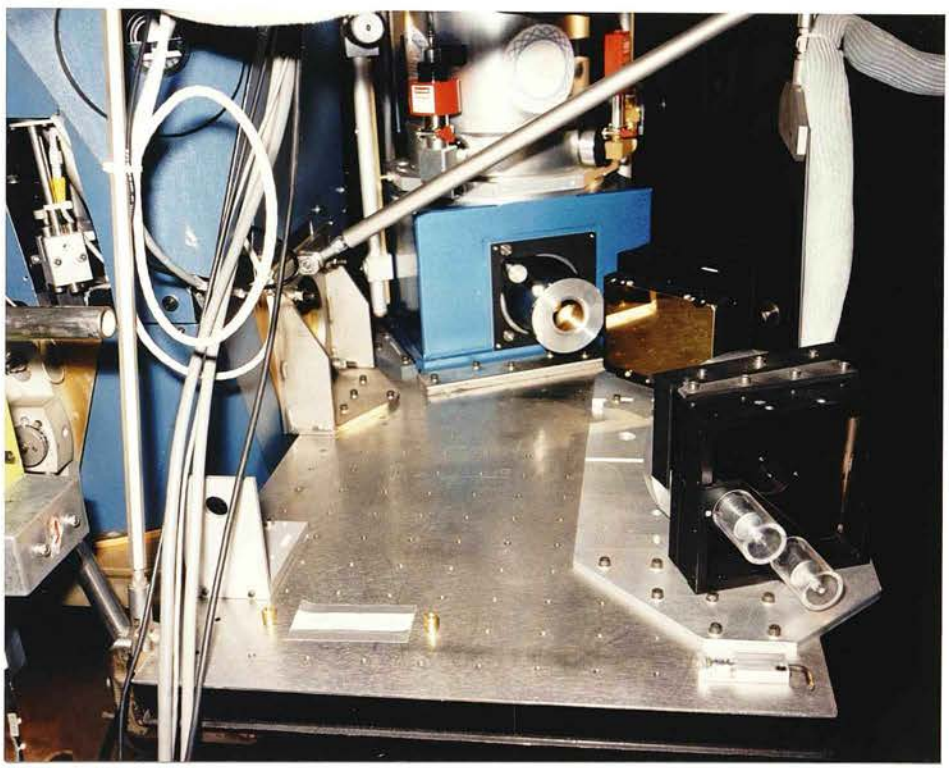


Figure 3.4 : Warm camera optics. In this view the IRCAM cryostat, plane mirror and collimating mirror are shown seated on the Instrument Mounting Platform. The f/36 beam emerges from the north port of the Instrument Support Unit (left).

field. Because of self obscuration the collimator mirror must be used 6° off–axis. A flat (plane) mirror in the collimator section "folds" the beam emerging from the focal plane in such a way as to minimise off–axis angles at the collimator which would otherwise introduce significant aberrations. The cryostat and the two gold–coated warm mirrors are mounted on an optical bench known as the Instrument Mounting Platform (IMP), which is rigidly supported from the Instrument Support Unit (ISU2) and the telescope primary mirror cell (see Figure 3.4).

3.3 Baffling

One of the requirements of the optical design was to form an image of the telescope aperture stop *inside* the cryostat to act as a cold stop (Lyot stop), this together with the fact that a camera has a well defined field of view at the detector, ensures that baffling via the intended optical path should be very good. This approach enabled the optical design to proceed without any further constraints imposed on it by the baffling; the baffles were literally built around the optical components. The baffles were designed by John Harris of the Royal Observatory, Edinburgh, and tested by the author.

It is instructive to consider the potential problem areas when no further precautions are taken. The collimator is the camera field stop and in the wide field (low–resolution) mode its image fills the detector array, and the warm highly emissive regions surrounding the collimator are imaged onto the regions surrounding the array. (In the narrow field mode this is less of a problem since the collimator overfills the array). Control of this flux becomes important beyond about 2 μm, where the thermal telescope background begins to dominate over the non–thermal, airglow emission from the sky. If the background flux imaged by the cold lens into the detector cavity (the cavity containing the detector at one end and the lens at the other) is not immediately absorbed it can be scattered onto the detector. This problem is accentuated by the fact that detector pixels have an unlimited field of view. The rays which can be directly imaged into the detector cavity are limited to relatively low inclinations by the solid angle subtended by the window at the Lyot stop (see Figure 3.5, ray AB). The precise inclination depends on where in the Lyot stop the window is viewed, but it is approximately $\theta = 15°$ and from Equation 2.17, the thermal background flux entering the detector cavity through

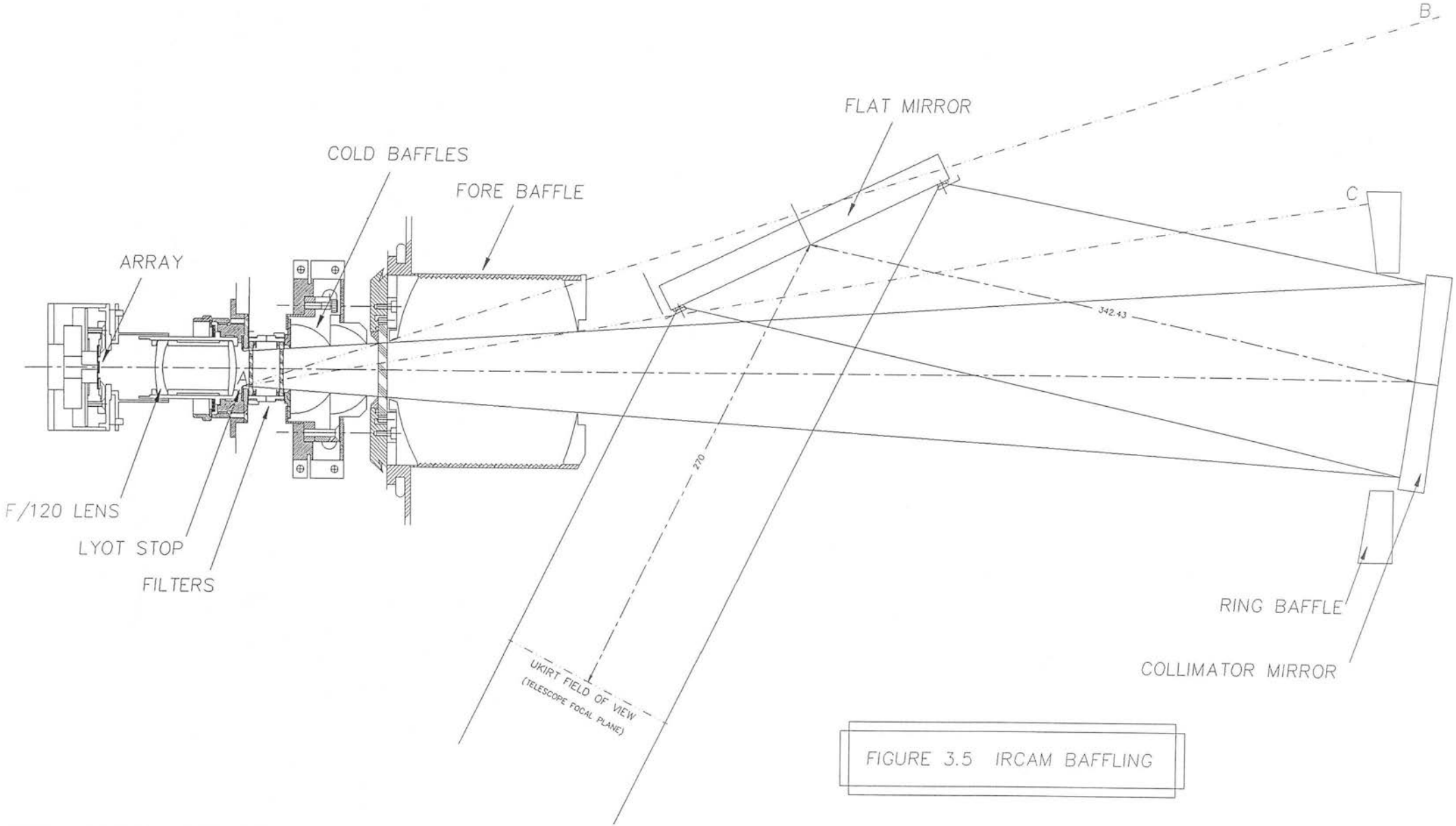


FIGURE 3.5 IRCAM BAFFLING

the filters in the low spatial resolution mode is given by :

$$F = \pi A_{Lyot} N(\lambda, 273) \Delta\lambda \left(\varepsilon_2 \left[\sin^2 \theta \right]_5^{15} + \varepsilon_1 \left[\sin^2 \theta \right]_0^5 \right) \tag{3.13}$$

where the collimator subtends a half-angle of $\theta = 5°$

ε_1 is the effective emissivity of the warm surfaces seen through the collimator ($\simeq$0.2)

ε_2 is the emissivity of the collimator surroundings ($\simeq$1)

Of this flux,

$$F = \pi A_{Lyot} N(\lambda, 273) \Delta\lambda \left(\varepsilon_1 \left[\sin^2 \theta \right]_0^5 \right) \tag{3.14}$$

is due to flux from the sky, and so the ratio of background to sky flux in the detector cavity is :

$$\frac{\varepsilon_2 \left[\sin^2 \theta \right]_5^{15} + \varepsilon_1 \left[\sin^2 \theta \right]_0^5}{\varepsilon_1 \left[\sin^2 \theta \right]_0^5} \simeq 40 \tag{3.15}$$

Consequently, if even a small part of this background were to be scattered onto the detector, it could severely degrade the S/N. As it is not possible to baffle the detector from this background by using another cold stop without occulting the optical beam, either the background flux must be absorbed in the cavity before scattering onto the detector, or the emissivity (ε_2) of the collimator surroundings must be minimised. Both these methods are employed. The warm surfaces surrounding the collimator are shielded with spherical mirrors having centres of curvature such that the detector cavity 'sees' the cold region immediately in front of the filters. In this situation the detector cavity only sees thermal emission from the mirrors themselves. For good mirrors ($\varepsilon_2 \rightsquigarrow$0.05), the ratio of background to sky flux in the low-resolution mode is reduced from $\sim$40 to $\sim$3. (Similar improvements apply to the medium and high spatial resolution modes.) In theory just one spherical mirror surrounding the field stop (the collimator) is required. However, because the optical axis is only 100mm above the optical bench, and due also to the position of the plane mirror, two mirrors are used. As shown in Figure 3.5, the first is located in the forebaffle. The outer radius of this mirror is defined by ray AB and the inner by the optical beam from the collimator. The ray AC defines the outer radius of the collimator mirror, while the inner radius is again set by the optical beam. From this construction it is clear that the longer the forebaffle the smaller the outer diameter of the collimator ring baffle needs to be, but that the length of the forebaffle is limited to that shown by the beam from the

telescope focal plane. Both warm baffle mirrors are gold coated. The detector cavity is blackened to absorb any background flux which does leak into it.

High and low inclination rays passing through the cryostat window but not through the Lyot stop, enter directly into the filter cavity, the cavity between the window and the filters. The ratio of this background flux to the sky flux in the low–resolution mode is :

$$\frac{\varepsilon_3 \left[\sin^2\theta\right]_{15}^{90} \Delta\lambda_3}{\varepsilon_1 \left[\sin^2\theta\right]_0^5 \Delta\lambda_1} \simeq 600 \frac{\Delta\lambda_3}{\Delta\lambda_1} \tag{3.16}$$

where ε_3 is the emissivity of the high inclination warm surroundings ($\simeq$1)
$\Delta\lambda_{1,3}$ is the bandwidth of the background and sky flux, respectively

The bandwidth of the sky flux is limited by the filters, however since there is a finite gap between the filter holders and the Lyot stop to allow the filter wheels to be rotated, and $\Delta\lambda_3 \gg \Delta\lambda_1$, significant amounts of out–of–band flux can leak into the detector cavity if care is not taken to minimise this gap. Even in the absence of leaks past the filters, multiple scattering of a high background in the filter cavity can cause significant leaks through the filters into the detector cavity, where it must be absorbed.

Ideally, high inclination rays entering the window can be excluded by extending the forebaffle mirror so that, together with a plane mirror containing a central hole for the window, it forms a hemispheric cavity. Such a geometry has the property that rays entering the cavity through the forebaffle aperture and not directly passing through the window into the filter cavity, can never do so, and the background flux entering the window is due solely to thermal emission from the low emissivity mirrors. As it was not practical to make such a large hemisphere (radius 150mm), a gold–coated baffle tube with a rectangular thread was used instead. Although far from ideal, the rectangular thread shares some of the flux rejection properties of the hemispheric baffle. However, it was possible to incorporate hemispheric baffles into the cavity immediately in front of the filters. If one hemispheric baffle were to be used it would require a radius of 45mm, which is too wide to fit. Two smaller baffles with the same flux rejection property are used instead. These baffles are made from brass and are heat treated to decrease (by oxidation) the reflectance of the surfaces, with the objective of not increasing the diffuse reflection, which would result if the surfaces were blackened by coating. Since the hemispheres are unlikely to be geometrically perfect, it is necessary to increase the

absorptive properties of the baffles, so that the intensity of rays which do leak through the plane mirror aperture, is significantly reduced.

At thermal wavelengths ($\lambda > 2.5\,\mu$m) the measured sky backgrounds are within about a factor of 2 of the expected levels (see Chapter 2.4.1), testament to the good performance of the baffling. Initially high backgrounds were traced to out–of–band photons leaking past the filters due to a grossly larger than specified gap between the filter holders and the Lyot stop.Low background operation was obtained when this gap was reduced to about 0.1mm.

3.4 Layout of the Camera Head

We have already shown that a good size for the cooled box containing the detector array and cold optics is 30cm x 30cm x 30cm. This is small enough to be cooled by a reasonably sized dewar yet big enough to contain the required cold optics and a large real–time choice of filters. The decision to use simple chromatic optics means that the focus is filter dependent and that a focusing mechanism is required. An estimate of the change in focus position to correct for longitudinal chromatic aberration can be found from the simple lens formula :

$$\frac{1}{f(\lambda)} = (n(\lambda) - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \quad \text{and} \quad \frac{df}{d\lambda} = \frac{f(\lambda)}{(n(\lambda) - 1)}\frac{dn}{d\lambda} \tag{3.17}$$

where $f(\lambda)$ is the wavelength dependent focal length

$n(\lambda)$ is the refractive index of the lens material

r_1,r_2 are the radii of curvature of the lens

Therefore the change in focal length due to a change in wavelength is given by :

$$\Delta f \simeq -\frac{f(\lambda)}{V} \quad \text{where} \quad V = \frac{n(\lambda) - 1}{\Delta n} \tag{3.18}$$

where V is known as the Abbe V–value

Δn is the change in refractive index over the wavelength interval $\Delta\lambda$

λ is the midpoint of the wavelength interval

Zinc Selenide (ZnSe) was chosen as the lens material because of its :

- high 1–5 μm transmittance when anti–reflection coated (total useful range 0.5–20 μm)
- high refractive index, n=2.4376 at λ=3 μm and T=290K (which from Equation 3.17 means larger radii of curvature, and therefore easier to make lenses)
- low dispersion $dn/d\lambda$ (V–values: 24 at 1–5 μm, 177 at 3–6 μm (McDonell and Klee 1984)

The greatest shift in longitudinal focus is suffered by the 120mm focal length lens.Over the wavelength range 1–5 μm the focus shift is 5mm. In order to focus an image it must first be defocused and from the optical spot diagrams the total depth of focus is found to be 0.5mm ($\leq \pm 0.25$mm). Therefore the total shift in focus required to cover the range 1–5 μm is :

$$\Delta f \simeq 5.0 + 0.5 = 5.5\text{mm}$$

Since the dispersion of ZnSe starts to increase quite rapidly shortward of 2 μm (Wolfe and Zissis 1978), and as the actual wavelength coverage is 1.2–4.8 μm, the fine focus is conservatively designed to move $\pm$3mm.

The IRCAM lens parameters are listed in Table 3.1. Serious consideration was initially given to mounting all three lenses in a turret mechanism driven by a gear mechanism and drive shaft. In this configuration the position of the detector assembly remains fixed and the desired image scale is selected in real–time by simply turning the turret until the correct lens is in the optical beam. However, to keep the same detector position for each (different focal length) lens, it is necessary for the distances between the front lens surface and the Lyot stop to be different. For the 120, 60 and 30mm focal length lenses, these distances are 0, 21 and 36mm respectively. Since the diameter of the collimated beam bundle increases with distance from the Lyot stop, the aperture of the each lens must be wide enough to accept the expanded beam. It needs to be widest in the case of the 30mm focal length lens, where a clear aperture of 24mm is required.

The ability to change image scale, to 'zoom in' on an interesting source, particularly in real time, is of great astronomical potential, and for this reason the lens turret

Table 3.1: **IRCAM Lens Parameters**

Lens	*Image Scale*	*Field of of View*	*Diameter (mm)*	*Clear Aper. (mm)*	BFL^a *(mm)*	
					295K	80K
30mm Doublet	2.4″/pixel	135″ [b]	30	24	49.93	46.18
60mm Meniscus	1.2″/pixel	74″ x 70″	30	24	60.8	61.3
120mm Telephoto	0.6″/pixel	39″ x 36″	30	24	79.9	82.1

[a] Back focal length, distance from the Lyot stop to the array, at λ=1.6 μm.

[b] Diameter of unvignetted field of view.

was included in the baseline design. This design therefore required 4 drive shafts, associated gear mechanisms and vacuum feed–throughs: one for the fine focus, one for the lens turret, and one for each of the two filter wheels. As the design proceeded it became clear that the lens turret mechanism was leading to some very tight spacing constraints within the camera head. Another concern was the possible deterioration in image quality, particularly in the low–resolution mode where, because of the expanded beam, more of the lens is used. The affects of flexure on the lens turret, and the more complicated baffling, were also considered problem areas. As it was felt that the additional complexity did not warrant the increased risk, the design was simplified by eliminating the lens turret. In the adopted approach, the lens is changed manually when the camera is warm. Although a lens change is a fairly trivial exercise, it does require the cryostat to be cold–cycled and exposes the detector array to air. Cold–cycling the hybrid array increases the risk of de–bonding the multiplexer from the InSb substrate, and since parts of the detector are hygroscopic, frequent exposure to air should be avoided. To minimise these dangers, changes of image scale are limited to twice per observing semester.

In the final design (see Figures 3.6, 3.7 and 3.8) an aluminium bulkhead contains the Lyot stop and filter wheel mounts. Each lens assembly is rigidly secured to the bulkhead by a brass retaining ring. Since the front lens surface is located at the Lyot stop, the detector position is wavelength dependent. Whenever the 120mm focal length lens is replaced with either the 60 or 30mm lens, the detector assembly must

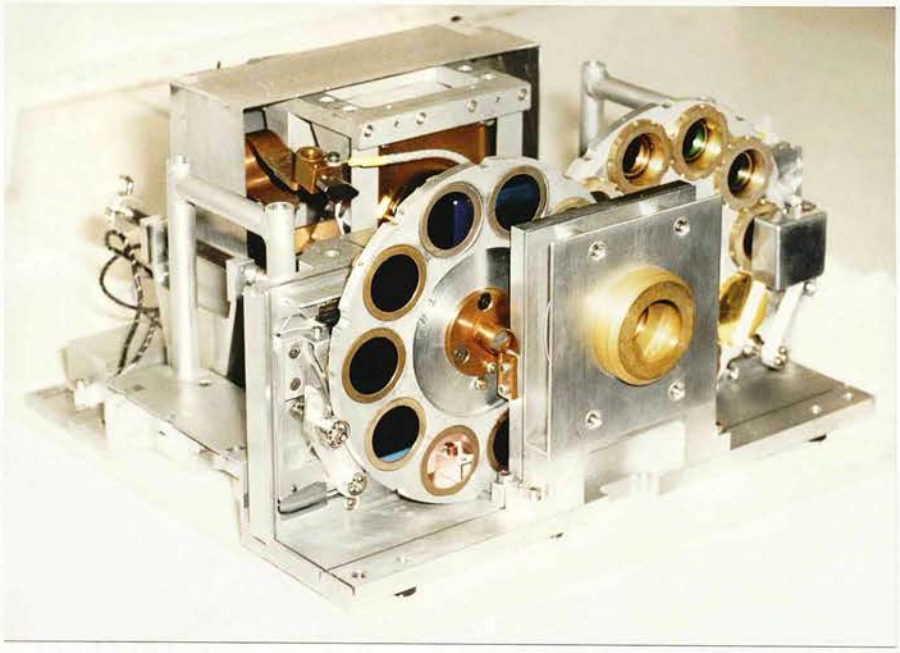


Figure 3.6 : Camera head, front view. The two filter wheels, the hemispherical brass baffles, and the detector and focus assembly are shown mounted to the cold table.

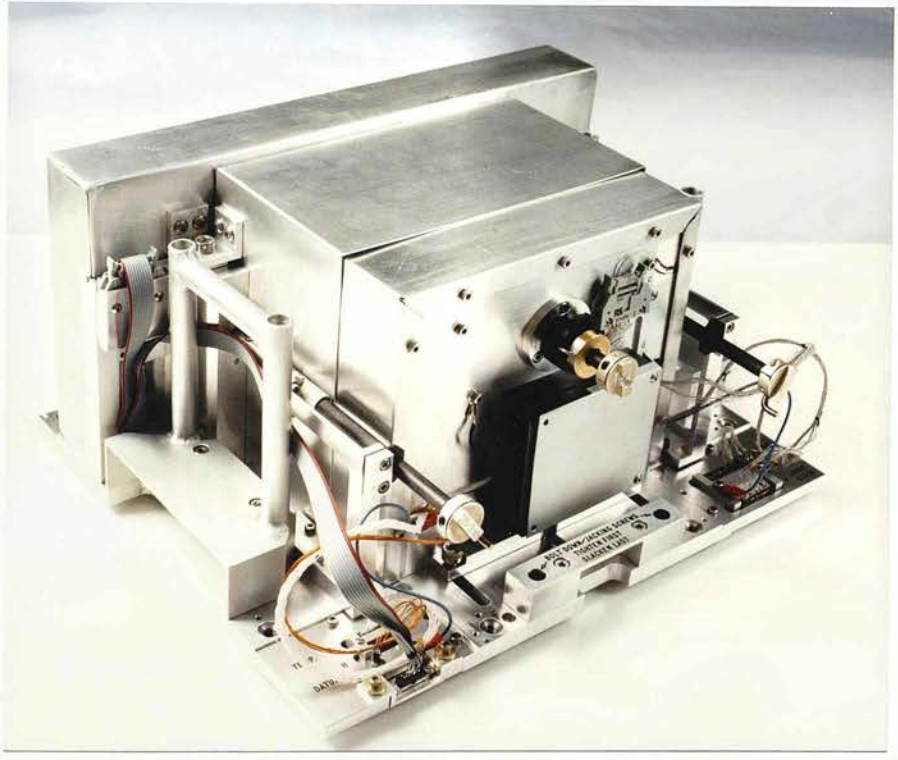


Figure 3.7 : Camera head, rear view. Baffle shields enclose the detector assembly and filter wheels, and shroud the detector from the 1–5 μm photons emitted by the drive shafts. Also visible is the pre–amp, which is mounted just below the focus drive shaft, and the two electrical mating points in each of the rear corners of the cold table.

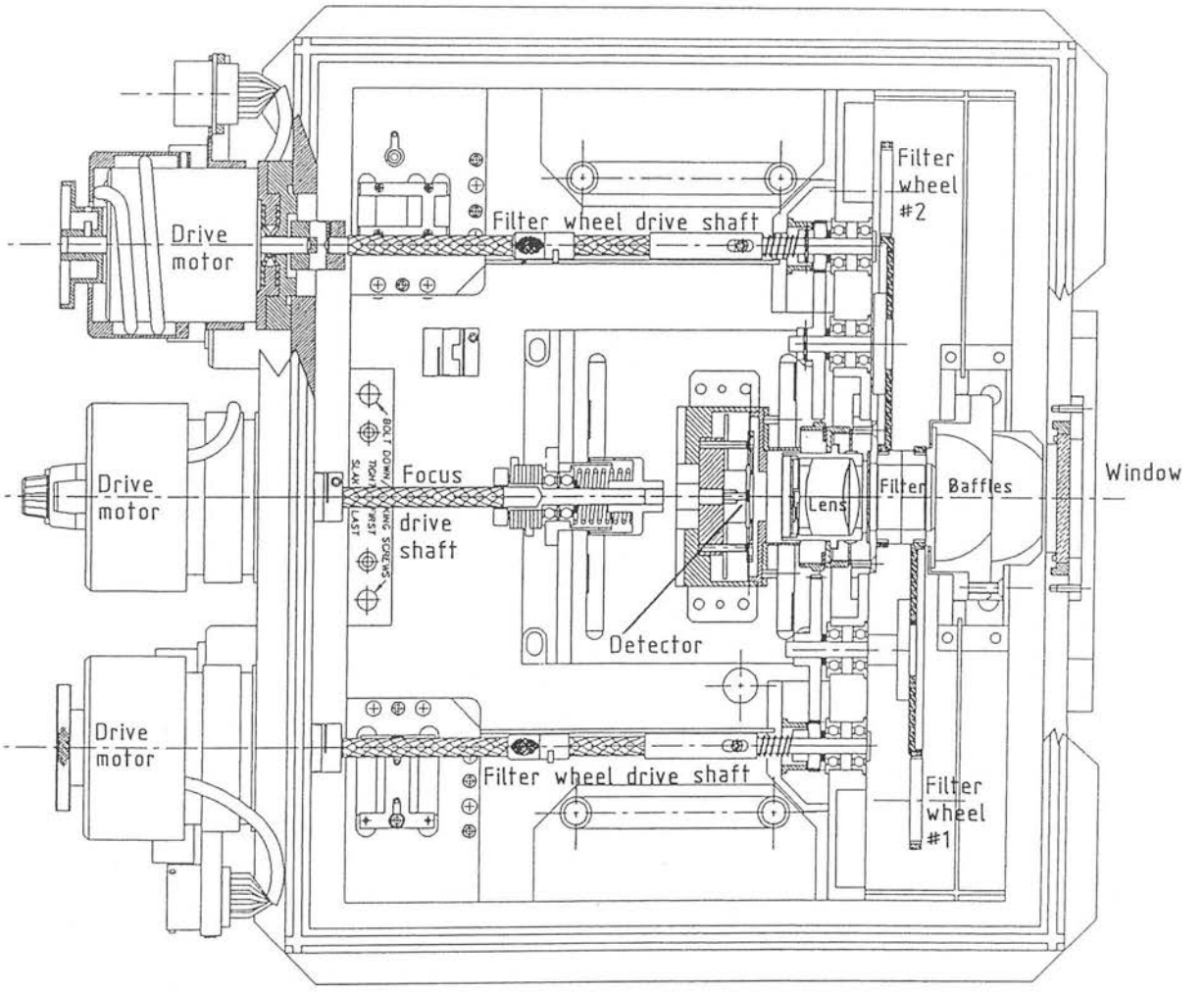


Figure 3.8 : Physical layout of the camera head. This configuration shows the f/30 wide angle doublet lens which is used in the low spatial resolution ($2.4\,\widehat{''}$ pixel^{-1}) imaging mode.

be moved towards the Lyot stop by 21 or 36mm respectively. The detector and fine focus assembly is manually repositioned by loosening the bolts which secure it to the optical table and then sliding it along the guiding slots in the table. Fine focusing is accomplished by a lead–screw mechanism in which the detector assembly is mounted on a four blade suspension system and the screw action moves the detector ±3mm in the direction normal to the axis of thin cross–section of the blades. (The cryo–mechanics of the Focus and Detector Assembly are fully discussed in Chapter 4.4.2.) The fine focus drive shaft incorporates a telescoping mechanism to compensate for the change in position. A nose–cone projecting from the detector assembly slides over the lens tube to provide excellent baffling. (To avoid unwanted heat loads into the detector assembly, the nose–cone (~35K) must not touch the lens tube (~80K).)

An important feature of the camera head is the mounting of all cold components on an optical table to provide ease of access and assembly. This enables the entire optical table to be removed without disturbing the relative alignment. The optical table is bolted to a rigid aluminium support ring inside the camera body and it can be removed by releasing 8 captive bolts and lifting up on the handles provided. When this is done the electrical mating contacts, which are mounted in the support ring, are automatically broken. To provide cooling, flexible copper braids connect the support ring to the LN_2 dewar. There are three components within the camera head which must be driven from outside : the two filter wheels and the fine focus mechanism. Each drive consists of identical 5–phase Berger Lahr stepping motors, rotary vacuum feed–throughs and drive shafts. The drive shafts are connected to the vacuum feed–throughs by easily demountable Oldham couplings, which also allow for thermal contraction. (In recent tests at ROE, stepping motors have been successfully operated at 80K, and so if space permits, there is now no need for warm vacuum feed–throughs in future instruments.) The filter wheels are driven through 3:1 gear wheels. Accurate positioning of the wheels is determined by detents on the rim of the wheel and by datum and detent microswitches. A discussion of the thermal implications of these drive mechanisms is given in Chapter 4.4.3.

The position of the detector assembly and by implication which lens is being used, is encoded by three microswitches. For each lens and filter combination the fine focus mechanism automatically drives the detector assembly to the position indicated in a software look–up table. The look–up table consists of a list of offsets from the broad

band K (2.2 μm) focus position, and so a K image must first be focused by moving the telescope secondary mirror.

Once the optical table is secured to the support platform and the drive shafts connected, a box–like cold radiation shield, with an aperture for the optical beam and slots for the drive shafts, is placed over it. The cold shield sinks the heat load radiated by the warm camera body, and which would otherwise be radiated onto the optical table. However, to shield the detector from the significant numbers of 1–5 μm photons emitted by the drive shafts, which unavoidably penetrate the cold shield box, the detector assembly and filter wheels are independently baffled (see Figure 3.7). From the outside the camera body appears as a rectangular box with an entrance window on its front face and all motor drives and electrical sockets emerging from its rear face; the side faces are unused. Since the cryostat is a vacuum chamber with a relatively large surface area, it has thick (~10mm) walls to resist deformation due to atmospheric pressure.

Chapter 4

Cryogenic Design

4.1 Introduction

The cryogenic camera IRCAM consists of a modified Oxford Instruments MD1810 dual vessel dewar connected to an optical instrument enclosure called the Camera Head. The camera head contains an optical table on which are mounted the cold optics and intermediate temperature detector assembly, and which is shrouded from the warm camera body by a LN_2 cooled radiation shield. The major requirements of the cryogenic design can be summarised as follows :

- Minimise thermal background photons from the camera head seen by the detector
- Maximise dewar hold–times
- Support the detector rigidly for imaging
- Operate the detector anywhere in the temperature range 10–50K (see Chapter 5)
- Minimise cool–down times
- Ensure ease of access and assembly

Thermal photons emitted by the optical enclosure and its contents are a potentially significant source of background photocurrent in the detector. We have already shown (see Chapter 2.4.3) that the optical table and its enclosure must be cooled to <85K for the background photocurrent they generate in the detector to be negligible compared to

the other sources of background photocurrent. This low temperature is accomplished by connecting the enclosure to the LN_2 vessel by a low thermal resistance interface while at the same time minimising the heat input to the cold enclosure, which flows through the interface into the cold enclosure. To avoid wasting expensive liquid cryogens and for the sake of operational flexibility, it is essential to maximise 'hold–times', that is the time between refilling the dewar vessels with cryogens. While hold–times of about 12 hours are sufficient to keep the instrument cold during an average observing session, it is of course highly desirable for the hold–times in to be comfortably in excess of one day. By optimising the dewar vessel capacities (4.5l of LN_2 and 4.0l of LHe) and minimising the heat inputs to the cold enclosure and to the detector assembly, LN_2 and LHe hold–times of $\sim$30 and $\sim$48 hours have been achieved. Imaging requires that the detector be rigidly supported. Meeting this requirement is made possible by mounting the detector rigidly on the optical table, which is itself mounted on the inherently stiff, but warm, camera body. To prevent significant heat inputs both these mountings must have high thermal resistances. Cool–down times, and therefore thermal cycling times, are also improved by a low thermal resistance interface and, additionally, by keeping the thermal mass of the cold components to a minimum. This is particularly important during testing of the instrument. Finally, it is important that the instrument design allows for ease of access and assembly.

In this chapter we will show in detail how these requirements have been realised. Of fundamental importance to the cryogenic design is an understanding of the heat flow mechanisms. This knowledge is used to construct a thermal model of the cryostat from which the cryogenic performance is optimised. The cryogenic mechanisms comprising the dewar/camera head interface, and the detector and focus assembly are considered in some detail. Both structures must satisfy the conflicting requirements of low flexure, low thermal resistance and compliance. The cryo–mechanics of the filter wheels and drive shafts are also investigated and in particular the contribution made to the cold enclosure thermal photon background by the drive shafts. Finally, the measured cryogenic performance is shown to compare favourably with that predicted.

4.2 Heat Inputs

4.2.1 Radiation

The net radiated heat transfer from a body at temperature T_2 onto a body at temperature T_1 and surface area A_1 is given by (see for example Wolfe and Zissis 1978) :

$$q = \sigma A_1 F_{2-1}(T_2^4 - T_1^4) \tag{4.1}$$

where q is the heat flow (usually in W)

σ is the Stefan–Boltzmann constant ($5.67\text{x}10^{-8}\text{Wm}^{-2}\text{K}^{-4}$)

F_{2-1} is the configuration factor

F_{2-1} is a factor which accounts for the emissivities of the surfaces and their geometrical relationship. (The geometrical and emissivity parts of this factor are sometimes separated into two factors (Chapman 1984)). Appendix A.1 discusses the configuration factor in some detail and derives the configuration factor for long co–axial cylinders and concentric spheres, which together with their approximations, are the two most commonly encountered geometries in cryostats. As can be seen from the derivation leading up to Equation A.22 the configuration factor for long co–axial cylinders or concentric spheres is given by :

$$F_{2-1} = \left[\frac{1}{1/\varepsilon_1 + (A_1/A_2)(1/\varepsilon_2 - 1}\right] \tag{4.2}$$

where A_2 and ε_2 are the surface area and emissivity of the outer body

A_1 and ε_1 are the surface area and emissivity of the inner body

The affect of introducing n 'floating' radiation shields, shields which effectively do not touch, is to reduce the radiated heat load on the inner body by a factor $(n + 1)$ (see Equation A.25). This form of Multiple Layer Insulation (MLI) is used between the warm outer cryostat body and the LN_2 cooled radiation shield. To have low emissivities metal surfaces need be only relatively thin (Equation A.12), and flexible aluminium foil and aluminised mylar wrapped radiation shields provide a low weight yet high packing density (large n) multiple layer insulation. In practice the shields are separated by low thermal conductivity netting and unavoidably rest on areas of the cryostat. Even though the shields are operated under high vacuum ($< 10^{-5}$mbar), there is inevitably

a heat load *conducted* through the MLI. These and other factors (see Shu *et al.* 1986, Shu 1987 and Shu *et al.* 1987) mean that the performance of MLI is usually much less than ideal. Nevertheless, MLI has been used to significantly reduce the cryostat heat loads (see Section 4.5).

4.2.2 Conduction

Solids

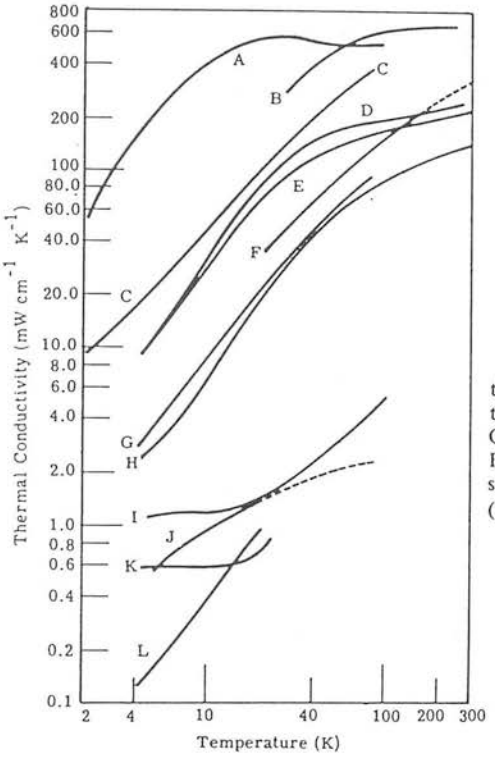
The fundamental equation for one–dimensional heat transfer by conduction is Fourier's Law (Chapman 1984) :

$$q = kA\frac{dT}{dx} \tag{4.3}$$

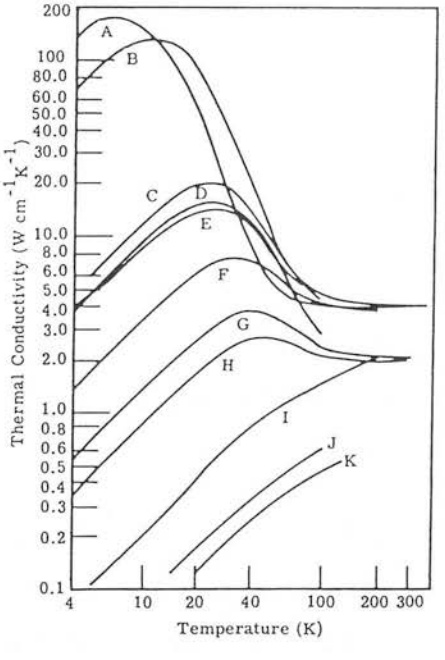
where q is the rate of heat transfer by conduction (usually in W)
k is the thermal conductivity of the material
A is the cross–sectional area through which the heat flows measured normal to the direction of heat flow
dT/dx is the temperature gradient at the section measured in the direction x of the heat flow

In this formulation edge effects are assumed to be negligible. The term "one–dimensional" is applied to a heat conduction problem when only one space coordinate is required to describe the temperature distribution within the body. To a very good approximation, all the conduction problems encountered in the cryogenic design can be considered to be one–dimensional. The basic unit of thermal conductivity is $\mathrm{Wm^{-1}K^{-1}}$ and the value of the thermal conductivity of solids at temperatures of 80K vary from $\sim 10^{-1}\mathrm{Wm^{-1}K^{-1}}$ for insulating materials (eg. G10 fibre-glass) to $\sim 10^{1}\mathrm{Wm^{-1}K^{-1}}$ for common constructional materials (eg. stainless steel, constantan wires) and up to $\sim 5\mathrm{x}10^{2}\mathrm{Wm^{-1}K^{-1}}$ for pure metals (eg. copper and aluminium).

Figure 4.1 shows the variation of thermal conductivity with temperature for a variety of solid materials. It can be seen that for most materials the thermal conductivity varies considerably with temperature. For a thermal conductor of length L and con-



Low-temperature, thermal conductivities of some solids with relatively low conductivities. A: 50-50 lead-tin solder; B: steel SAE 1020; C: berylium copper; D: constantan; E: monel; F: silicon bronze; G: inconel; H: type 347 stainless steel; I: fused quartz; J: polytetrafluorochylene (Teflon); K: polymethacrylate; L: nylon



Low-temperature, thermal conductivities of some metals with relatively high conductivities. A: silver, 99.999% pure; B: high purity copper; C: coalesced copper; D: copper; E: aluminum, single crystal; F: free-machining tellurium copper; G: aluminum, 1100; H: aluminum, 6063-T5; I: copper, phosphodeoxidized; J: aluminum, 2024-T4; K: free-machining leaded brass

Figure 4.1 : Variation of thermal conductivity with temperature (from Wolfe and Zissis 1978, original source Scott 1959).

stant cross–section A Equation 4.3 can be rewritten in the form :

$$q = \frac{A}{L}\int_{T_1}^{T_2} k dT \tag{4.4}$$

The integral of k with respect to temperature is known as a Thermal Conductivity Integral and it is tabulated for a number of materials (Goodall 1970). Alternatively the variation can be allowed for by curve fitting. The physical mechanism of heat conduction in solids and the reasons for the variation of thermal conductivity with temperature will not be discussed here. For a fuller discussion see Barron (1985).

A strict solution of the heat flow through a thermal conductor must include the heat lost (or gained) from the surface of the conductor. An important case in point is the heat flowing through the drive shafts which penetrate the cold optical enclosure. In this situation the heat flowing through the shafts is *not* constant, and the difference is equal to the heat radiated from the surface of the conductor to the cold enclosure :

$$kA\frac{d^2T}{dx^2} = \sigma\varepsilon_s(2\pi r)\left[T^4 - T_e^4\right] \tag{4.5}$$

where $d\overline{T}/dx$ is the temperature gradient in the shaft
ε_s is the emissivity of the surface of the shaft (cavity approximation)
r is the radius of the shaft
T_e is the temperature of the enclosure (80K)

The wall of the hollow drive shaft is thin enough for this to remain a one–dimensional heat transfer problem. Equation 4.5 is a second order first degree non–linear differential equation and is solved numerically for T as a function of distance along the shaft x. The solution when applied to the drive shafts (see Section 4.4.3) gives the total heat input to the cold enclosure and the number of 1–5 μm photons from which the detector must be shielded.

The temperature drop across a thermal conductor due to heat flowing through it is analogous to the voltage drop across an electrical conductor (or resistor) due to an electrical current. On the basis of Equation 4.3, we can define a thermal resistance R by :

$$R = \frac{\Delta T}{q} = \frac{1}{k}\frac{\Delta x}{A} \tag{4.6}$$

In a thermal circuit it is easy to show that for a one–dimensional heat flow, the total

thermal resistance R_T, of n thermal conductors connected in series is simply :

$$R_T = R_1 + R_2 + ... + R_n = \sum_{i=1}^{n} \frac{1}{k_i} \frac{\Delta x_n}{A_n} \tag{4.7}$$

In practice, when thermal conductors are connected together (eg. two metal plates bolted together), an additional temperature drop will occur due to imperfect contact at the interface. By definition, the thermal conduct resistance is given by :

$$R_{in} = \frac{\Delta T_{in}}{A_{in} q_{in}} = \frac{1}{A_{in} h_{in}} \tag{4.8}$$

where q_{in} is the heat flowing through the interface per unit area

A_{in} is the apparent contact area of the interface

ΔT_{in} is the temperature drop across the interface

h_{in} is defined as the thermal interface conductance

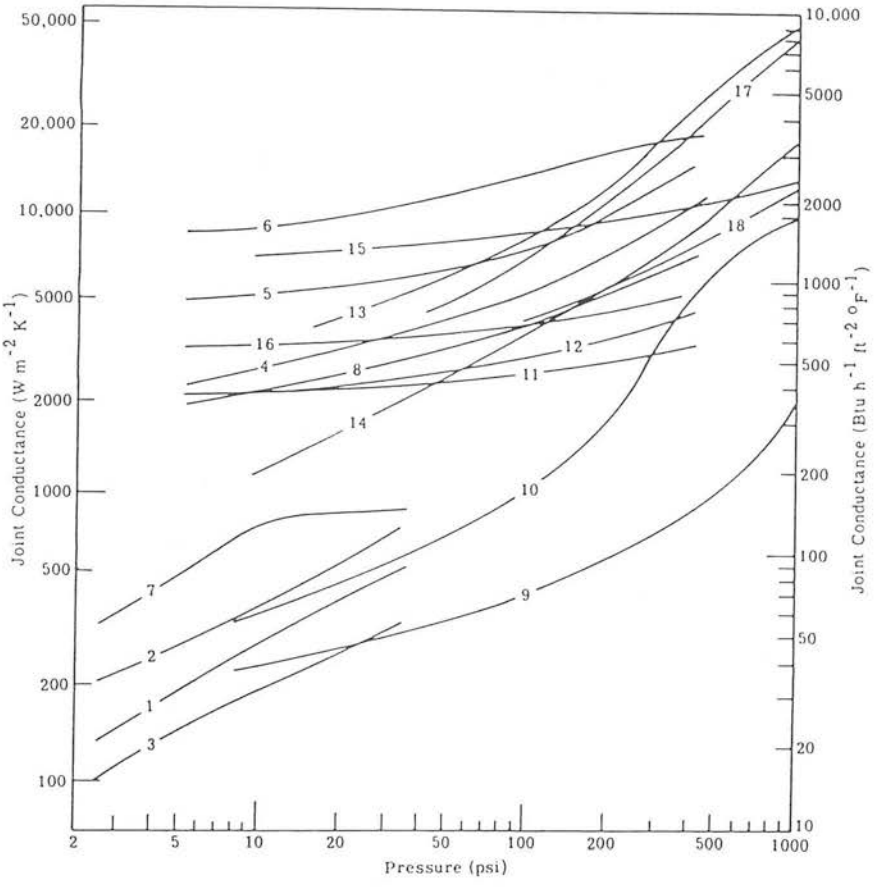
Heat transfer through the interface takes place by the combined mechanisms of conductance across true contact areas, conduction across entrapped interstitial fluids, and radiation across interstitial gaps. The resulting overall conductance of the interface is a function of the materials in contact (conductivity, surface finish, flatness and hardness), the contact pressure, the mean temperature of the interface, the nature of the interstitial fluid (liquid,gas,vacuum), and the presence of oxide films or interstitial shim materials (Rohsenow and Hartnet, 1973). The apparent contact area A_{in} is just the surface area of the interface (which is not necessarily in contact). The actual material–to–material contact area A_c is contained within the definition of the interface conductance and so $h_{in} \propto A_c$. From simple elastic deformation theory it follows that :

$$h_{in} \propto \frac{1}{E} \left(\frac{F}{A_{in}} \right) \tag{4.9}$$

where E is Young's modulus (Nm^{-2})

F/A_{in} is the tensile stress or pressure due to tensile force F

Figure 4.2 is a plot of the thermal interface conductance against contact pressure for a variety of materials and conditions. As predicted, the conductance generally increases with pressure. Significant increases of interface conductance can also be obtained by the application of interstitial fluids such as silicone high vacuum grease (a factor of 5 on brass surfaces, see Clausing and Chao 1965), and interface shim materials which are softer than the interfaces. These improvements are due to an effective increase in the area of contact.



Curve	*Material Pair*	*rms Surface Finish* (μin.)	*Gap Material*	*Mean Contact Temp.* (°F)
1	Aluminum (2024-T3)	48-65	Vacuum (10^{-4} mm Hg)	110
2	Aluminum (2024-T3)	8-18	Vacuum (10^{-4} mm Hg)	110
3	Aluminum (2024-T3)	6-8 (not flat)	Vacuum (10^{-4} mm Hg)	110
4	Aluminum (75S-T6)	120	Air	200
5	Aluminum (75S-T6)	65	Air	200
6	Aluminum (75S-T6)	10	Air	200
7	Aluminum (2024-T3)	6-8 (not flat)	Lead foil (0.008 in.)	110
8	Aluminum (75S-T6)	120	Brass foil (0.001 in.)	200
9	Stainless (304)	42-60	Vacuum (10^{-4} mm Hg)	85
10	Stainless (304)	10-15	Vacuum (10^{-4} mm Hg)	85
11	Stainless (416)	100	Air	200
12	Stainless (416)	100	Brass foil (0.001 in.)	200
13	Magnesium (AZ-31B)	50-60 (oxidized)	Vacuum (10^{-4} mm Hg)	85
14	Magnesium (AZ-31B)	8-16 (oxidized)	Vacuum (10^{-4} mm Hg)	85
15	Copper (OFHC)	7-9	Vacuum (10^{-4} mm Hg)	115
16	Stainless/aluminum	30/65	Air	200
17	Iron/aluminum	–	Air	80
18	Tungsten/graphite	–	Air	270

Figure 4.2 : Thermal contact conductance versus contact pressure (from Wolfe and Zissis 1978, original source Rohsenhow and Hartnett 1973). Examples of measured contact conductance data h_i as a function of contact pressure for the interface conditions listed in the accompanying table.

The cold optical table is cooled by attaching it to the LN_2 vessel by several thermal conductors and bolted interfaces connected together in series. Since it is important to keep the thermal resistance of the connecting path low, this understanding of the factors affecting the interface conductances is applied to the design of the dewar/camera head interface (see Sections 4.3 and 4.4.1).

Gases

The conduction of heat by a gas is discussed in detail in Chapter 5.2. When the gaseous mean free path is much greater than the dimensions of the apparatus (eg. the distance between the 300K camera body and the 80K LN_2 cooled radiation shield in the case of the cryostat) the conducted heat flow is proportional to the gas pressure. Writing Equation 5.3 explicitly for helium and nitrogen gas, assuming an overall accommodation coefficient of $a_0 = 0.25$ (ie $a_1 = a_2 = 0.5$), and a geometry corresponding to co–axial cylinders ($A = A_1 = A_2$), we obtain :

$$\text{Helium} \quad \frac{H}{A} = 5.3\ P(T_2 - T_1) \tag{4.10}$$

$$\text{Nitrogen} \quad \frac{H}{A} = 30\ P(T_2 - T_1) \tag{4.11}$$

where H/A is the heat flow in Wm^{-2}

T_1, T_2 are the absolute surface temperatures

P is the gas pressure in mbar measured at the room temperature gauge

When the inner vessel of the dewar contains LHe, pressures of $\sim 10^{-5}$ mbar are routinely obtained. Applying Equation 4.10 (assuming all the other gases are condensed onto the LHe cooled surfaces) this results in a heat load due to gas conduction of ~ 0.01W to the LN_2 vessel ($A = 0.7\mathrm{m}^2$, $T_2 - T_1 = 210$K) and ~ 0.005W to the LHe vessel ($A = 0.15\mathrm{m}^2$, $T_2 - T_1 = 70$K), both of which are small compared to the other expected heat loads (see Section 4.5). It is most important to ensure that the cryostat contains a good vacuum *before* attempting to fill it with liquid cryogens, otherwise gases maybe adsorbed ('cryopumped') onto the surfaces of the vacuum insulation. This increases the emissivity and results in a greater radiated heat load.

4.2.3 Convection (vapour cooling)

Since the cryostat is operated under very low vacuum pressures, convection is absent. However, forced convection is used in the dewar to remove heat conducted down the vapour vent tubes. The Oxford Instruments MD1810 dewar uses the cooling power of helium vapour venting from the inner vessel, which would otherwise be wasted, to cool a shield surrounding the liquid helium vessel. This Helium Vapour Shield is cooled to a temperature less than the LN_2 cooled shield, and consequently reduces the heat load on the LHe vessel. For a perfect gas or for an expansion done at constant pressure, the heat absorbed by a material is given (see for example Adkins 1975) by :

$$dH = mC_p dT \tag{4.12}$$

where dH is the change in enthalpy (J)
m is the mass of the material (Kg)
dT is the change in temperature (K)
C_p is the specific heat at constant pressure ($Jkg^{-1}K^{-1}$)

As indicated in the following analysis the effectiveness of the vapour shields in reducing the heat load depends upon how much of the heat load (base load B) is unrelated to the cold shield temperatures (eg. the heat conducted into the LHe vessel from the detector and its support structure).

From Equation 4.1, heat radiated onto the vapour shield from the LN_2 cold shield can be written as :

$$Q_{2-V} = \sigma F_{2-V} A_V (T_2^4 - T_V^4) \tag{4.13}$$

where T_2 is the temperature of the LN_2 cold shield (80K)
T_V is the temperature of the helium vapour shield
A_V is the surface area of the helium vapour shield
F_{2-V} is the configuration factor of the LN_2 cold shield as seen by the helium vapour shield

The heat flowing into the LHe vessel is due to the conducted base load B and the radiated heat load from the vapour shield :

$$Q_{\rightarrow 1} = \sigma F_{V-1} A_1 (T_V^4 - T_1^4) + B = \dot{m}h \tag{4.14}$$

where T_1 is the temperature of the LHe vessel (4K)

A_1 is the surface area of the LHe vessel

F_{V-1} is the configuration factor of the vapour shield seen by the LHe vessel

$\dot{m}$ is the mass flow rate of the boil–off vapour

h is the heat of vapourisation of LHe (Jkg^{-1})

Heat flowing out of the vapour shield due to radiation and the cooling effect of the venting helium vapour (the change in enthalpy, see Equation 4.12) is given by :

$$\begin{aligned} Q_{V\rightarrow} &= \sigma F_{1-V} A_V (T_V^4 - T_1^4) + \eta C_p \dot{m} (T_V - T_1) \\ &= \sigma F_{1-V} A_V (T_V^4 - T_1^4) + \left[\sigma F_{V-1} A_1 (T_V^4 - T_1^4) + B\right] (T_V - T_1) \eta \frac{C_p}{h} \end{aligned} \quad (4.15)$$

where F_{1-V} is the configuration factor of the LHe vessel seen by the vapour shield

η is the vapour cooling efficiency and allows for the fact that the vapour never reaches the temperature of the vapour shield ($\eta < 1$)

From conservation of energy :

$$Q_{2-V} = Q_{V\rightarrow} \quad (4.16)$$

Equations 4.13, 4.14 and 4.15 can be solved for T_V. The ratio of the heat load on the LHe vessel with (Q) and without (Q_o) a vapour cooled shield is given by :

$$\frac{Q}{Q_o} = \frac{\sigma F_{V-1} A_1 (T_V^4 - T_1^4) + B}{\sigma F_{2-1} A_1 (80^4 - 4^4) + B} \quad (4.17)$$

Figure 4.3 shows a plot of Q/Q_o against the base heat load B for different vapour cooling efficiencies η, (all the configuration factors are estimated to be ~0.025). A simple heat exchanger (see Figure 2.11) is used to force the venting vapour into closer contact with the wall of the vent tube at the location of the root of the helium vapour shield in an attempt to improve the cooling efficiency. Measurements of the vapour shield temperature as a function of B indicate $\eta \sim 0.05$. However, for the base loads experienced in the IRCAM cryostat ($B \simeq 50$mW), it can be seen from Figure 4.3 that Q/Q_o is relatively insensitive to η and that the vapour shield results in a reduction in heat load of about 13%. It is also apparent that even if the shield is only radiatively cooled ($\eta = 0$) the heat load is still reduced by about 6%. Vapour shield cooling becomes very efficient at very small base loads and from Equation 4.15 it can be seen

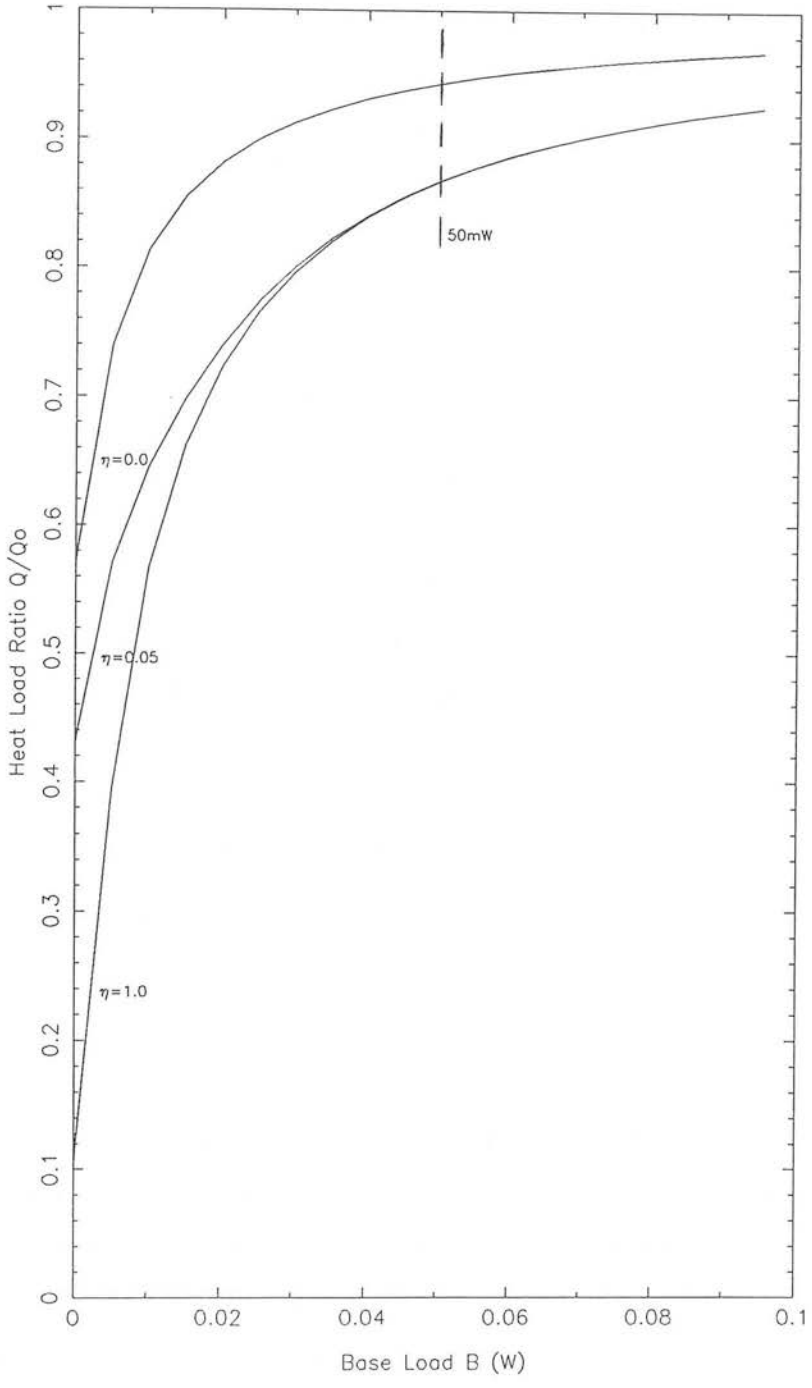


Figure 4.3 : Effectiveness of the helium vapour cooled shield. The ratio of the heat load on the liquid helium vessel with (Q) and without (Qo) a helium vapour cooled shield, is plotted against the base heat load B, for different vapour cooling efficiencies η (best estimate η=0.05), using $F_{V-1} = 0.025$, A_V =0.2m^2 and A_1 =0.15m^2. B is the total conducted heat load on the liquid helium vessel (ie. any heat input that does not depend on the temperature of the radiation shields); for the IRCAM cryostat B is typically 50mW.

Table 4.1: **Vapour Cooling Capacities**

Fluid	C_p/h (K^{-1})
Helium	0.25
Neon	0.013
Nitrogen	0.005

that the amount of vapour cooling is proportional C_p/h, and from Table 4.1 it is clear that vapour shields are most usefully employed in LHe dewars.

4.3 Cryostat Thermal Model

To optimise the performance of the IRCAM cryostat a thermal model was developed (see Figure 4.4). By modifying the various parameters in the model the sensitivity of the cryogenic performance to these changes was investigated and potential improvement areas identified. The objective of the optimisation was to :

- Obtain a low enclosure temperature (<85K)
- Maximise hold–times
- Minimise cooling times

This was achieved by identifying the major sources of heat input and precisely where in the cryostat they occur in relation to the cooling path. As illustrated in Figure 4.4 the thermal model is represented by a number of temperature nodes linked together by thermal resistances and radiation links. An accurate model of the cryostat requires a large number of temperature nodes, however, in practice the model is limited by knowledge of the parameters (eg. thermal conductivities, dimensions, contact resistances etc), consequently a thermal model consisting of relatively few temperature nodes can still be very useful. In particular the effective emissivity (ie. the emissivity part of the configuration factor) of the MLI, is notoriously difficult to predict accurately.

The cooling time depends upon the change in enthalpy of the mass to be cooled (see Equation 4.12), and the thermal resistance (or thermal conductance) of the path

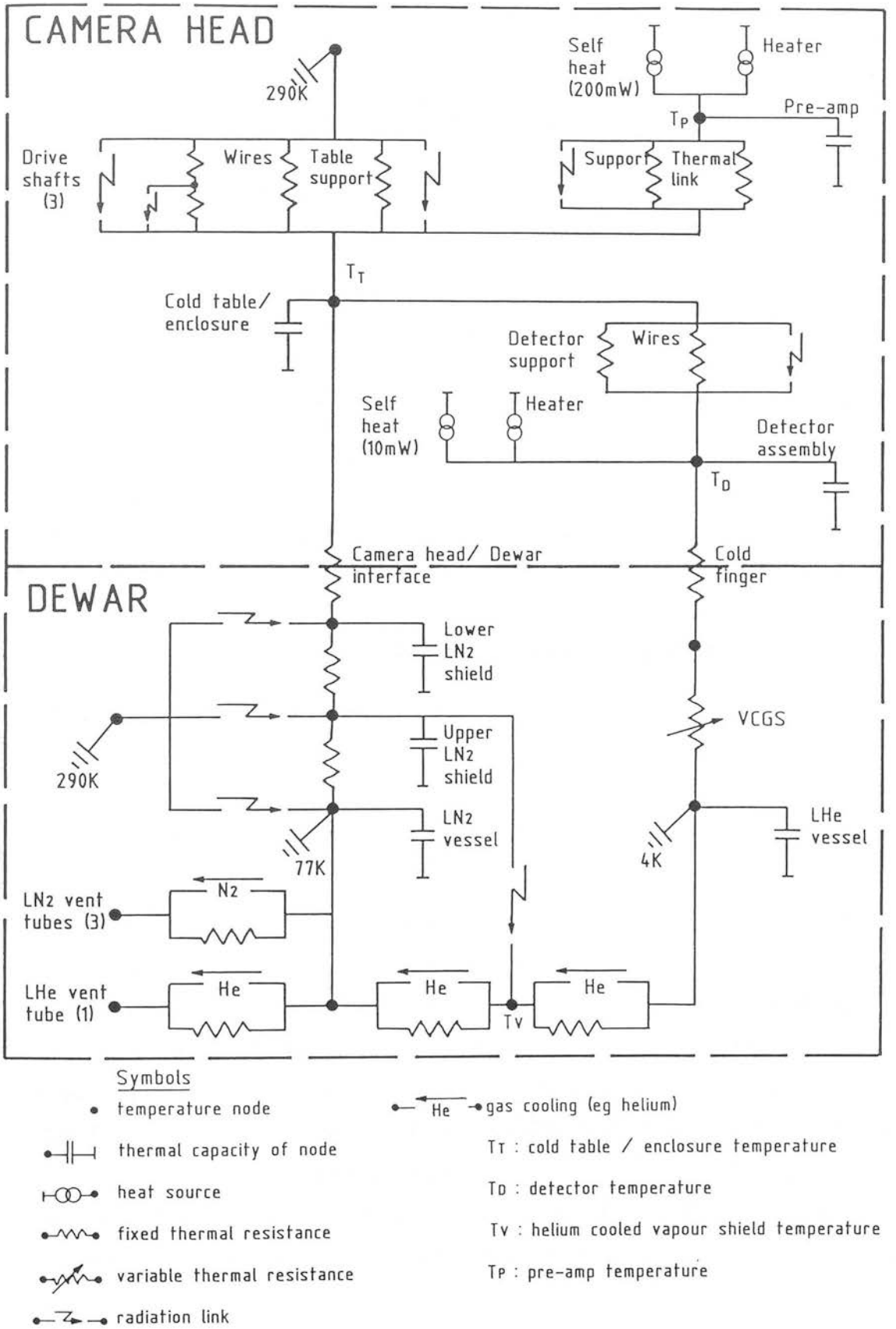


Figure 4.4 : Cryostat thermal model.

connecting the mass to the source of cooling :

$$\text{Cooling time } (T_i \text{ to } T_{i+1}) = \frac{mC_p(T_i)[T_i - T_{i+1}]}{(A/L)k(T_i)[T_i - T_{sink}] - Q(T_i)} \quad \text{where } T_i > T_{i+1} \tag{4.18}$$

and where T_i is the initial temperature

T_{i+1} is the temperature to be cooled to

T_{sink} is the temperature of the sink (eg 77K for the LN_2)

$Q(T_i)$ is the heat input to the mass at temperature T_i

For small cooling intervals, $k(T_i) \simeq k(T_{i+1})$ and $C_p(T_i) \simeq C_p(T_{i+1})$. The total cooling time is just the summation of Equation 4.18 over the individual cooling intervals. The cooling time becomes infinite (temperature equilibrium established) when the cooling capacity ($(A/L)k\Delta T$) becomes equal to the heat input Q. Equation 4.18 separates into two conditions :

1. $\left(\frac{A}{L}\right) k[T_{eq} - T_{sink}] = Q(T_{eq})$: Thermal equilibrium (*steady state*)
2. $\left(\frac{A}{L}\right) k[T_i - T_{sink}] > Q(T_i)$: Cooling (*transient state*)

The first condition can be expressed as :

$$[T_{eq} - T_{sink}] = Q(T_{eq}) \left(\frac{L}{A}\frac{1}{k}\right)$$

and so to obtain a low equilibrium temperature we must minimise the heat input ($Q(T_{eq})$) and the thermal resistance ($\frac{L}{A}\frac{1}{k}$). These are the same measures that must be taken to reduce the cooling time. The cooling time is proportional to the thermal mass (mC_p) of the material to be cooled. For ease of construction , the components of the camera head (filter wheels, focus assembly, optical table etc) are made from aluminium alloy. Obviously, for fast cooling, and low flexure of the cold enclosure, it is desirable to minimise the mass of the cold enclosure. However, the cold enclosure components must maintain internal rigidity, and so the mass must be optimised. Similar arguments as regards cooling and mass apply to the detector assembly.

4.3.1 Camera Head

The sources of heat input to the cold enclosure are :

- Radiation from the warm camera body. Due to the simple geometry, this is modelled as a single radiation link.

- Drive shafts. Three fibre–glass drive shafts penetrate the cold enclosure and the resulting heat load is due to conduction and radiation off the shafts. Although only two temperature nodes are drawn in the thermal circuit, the actual thermal model contains over 100 (see Section 4.4.3). Cold baffles shield the detector from the drive shafts.

- Cold table suspension. The cold table is rigidly mounted from the warm camera body by 4 fibre–glass thermal isolation pillars (each connected to an aluminium alloy blade to allow for thermal contraction). Since these 4 structures are mounted outside the cold enclosure, the radiated heat load can be ignored, and so the suspension system can be considered as one thermal resistance.

- Wires. About 150 electrical wires interface with the cold table electrical connectors. Again, since they do not penetrate the cold enclosure, the wires can be considered as a single thermal resistance.

- Pre–amplifier. For efficient electrical performance the detector pre–amp is positioned near to the detector (short wires) and operated at a controlled temperature of 120K. It is rigidly mounted on the cold table by a delring thermal isolator. By carefully trimming the length of a copper wire thermal link, the 200mW of self heating in the pre–amp gives rise to an equilibrium temperature of of just below 120K. Fine temperature control is accomplished by feedback control of an electrical resistance heater. Although the thermal affects of the heat load radiated by the pre–amp are negligible, cold baffles shroud the detector from the significant numbers of 1–5 μm photons it emits.

- Window. Due to efficient baffling (see Chapter 3.3), the heat which radiates into the cold enclosure is insignificant compared to the other sources.

- Gas Conduction. When the cryostat is operated at pressures below 10^{-5} mbar gas conduction is very small.

All the conducted heat loads are minimised by using low thermal conductivity materials (such as fibre–glass for support structures ,and constantan wires), and small A/L ratios where possible.

There are two heat flow paths into the dewar from the camera head. The first provides a cooling path to the LN_2 temperature sink for the cold enclosure. All the heat inputs to the cold enclosure flow into the LN_2 vessel through the dewar/camera head interface resistance which is connected in series with the LN_2 cold shield resistance. To achieve a low cold enclosure temperature and a short cooling time, it is vital to minimise these resistances. The dewar/camera head interface is drawn here as a single thermal resistance, but it is made up of several resistances in series. The LN_2 cold shield as well as providing a low thermal resistance connection to the cold enclosure, must also shroud the helium vapour cooled shield from room temperature radiation. It is modelled as two temperature nodes into which heat flows both by conduction and radiation. The second heat flow path connects the detector assembly (or cold stage) to the LHe temperature sink through the variable conductance gas switch. To minimise the heat flowing into the LHe vessel, the detector assembly is mounted from the cold table by high resistance fibre-glass pillars and the heat radiated onto the detector assembly is minimised by a low emissivity gold coating. The cold finger which connects the detector assembly to the gas switch must have a low thermal resistance since the resistance between the detector and the LHe cold–face needs to be controlled solely by the gas switch. Figure 4.5 is a photograph of the dewar section of the dewar/camera head interface, together with the copper cold finger which projects upward from the gas switch mounted on the cold–face. The ring assembly which bolts the cold table, is connected to the LN_2 cold shield by 40 flexible copper braids.

4.3.2 Dewar

The major sources of heat input to the dewar are :

LN_2 Vessel

- Camera head heat load
- Radiation from the warm camera body This is shown as 3 separate radiation links since it affects the temperature of the LN_2 cold shield. However, the temperature gradient in the cold shield has little affect upon the total heat load radiated from the warm dewar wall onto the cold shield, and so from this point of view it

can be considered as a single radiation link.

- Vent tubes (Three nitrogen tubes and one helium tube.) These are considered as temperature nodes connected to the LN_2 vessel by thermal resistances (stainless steel tubes), with the nodes cooled by the venting vapour. (The enthalpy change of helium vapour is 384 Jgm^{-1} from 4–77K and 1542 Jgm^{-1} from 77–273K, while for nitrogen vapour it is only 234 Jgm^{-1} from 77–273K).

- Gas conduction This is very small at pressures $< 10^{-5}$mbar.

LHe Vessel

- Detector assembly heat loads Detector power dissipation, conduction, radiation and gas switch power dissipation.

- Radiation from the helium vapour cooled shield

- Vent tube (One helium tube.) This is considered as a temperature node, nominally at the same temperature as the vapour shield, connected to the LHe vessel by a thermal resistance (a stainless steel tube) and which is cooled by the venting helium vapour.

- Gas conduction This is very small at pressures $< 10^{-5}$mbar.

The Oxford Instruments MD1810 Dewar was bought as a standard item and so it was not planned to optimise its design. Using the specified performance of the dewar, an estimate made of the effective emissivity of the of its multiple layer insulation (F_{2-1}=0.025 in Equation 4.1) was made, and this value was used in the thermal model of the cryostat for the MLI in the camera head. The resulting model predicted an equilibrium cold enclosure temperature of about 90K, 5K in excess of the specification. This was due to the higher than desired thermal resistance of the aluminium alloy LN_2 cold shield, which had been designed as a temperature shroud, not as a heat sink path. The problem was solved by using a cold shield with a thicker wall (see Section 4.4.1).

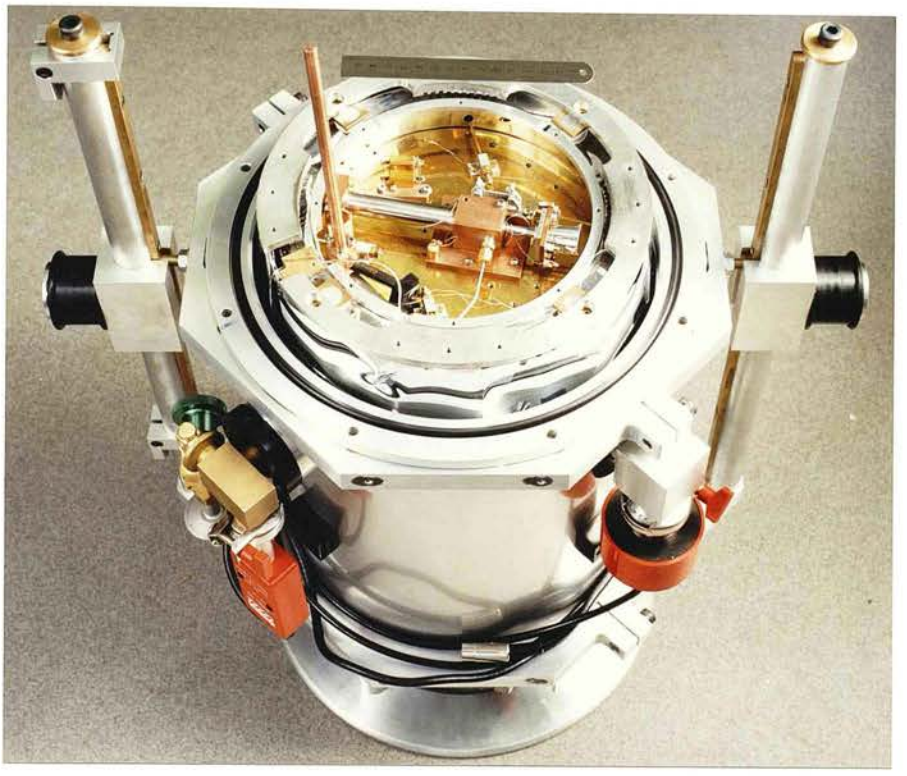


Figure 4.5 : Dewar section of the dewar/camera head interface. Visible is the gas switch mounted on the LHe cold face, the ring assembly which is connected to the LN_2 cold shield and which bolts to the cold table in the camera head, and the 'O' ring which forms a vacuum seal with the camera head case.

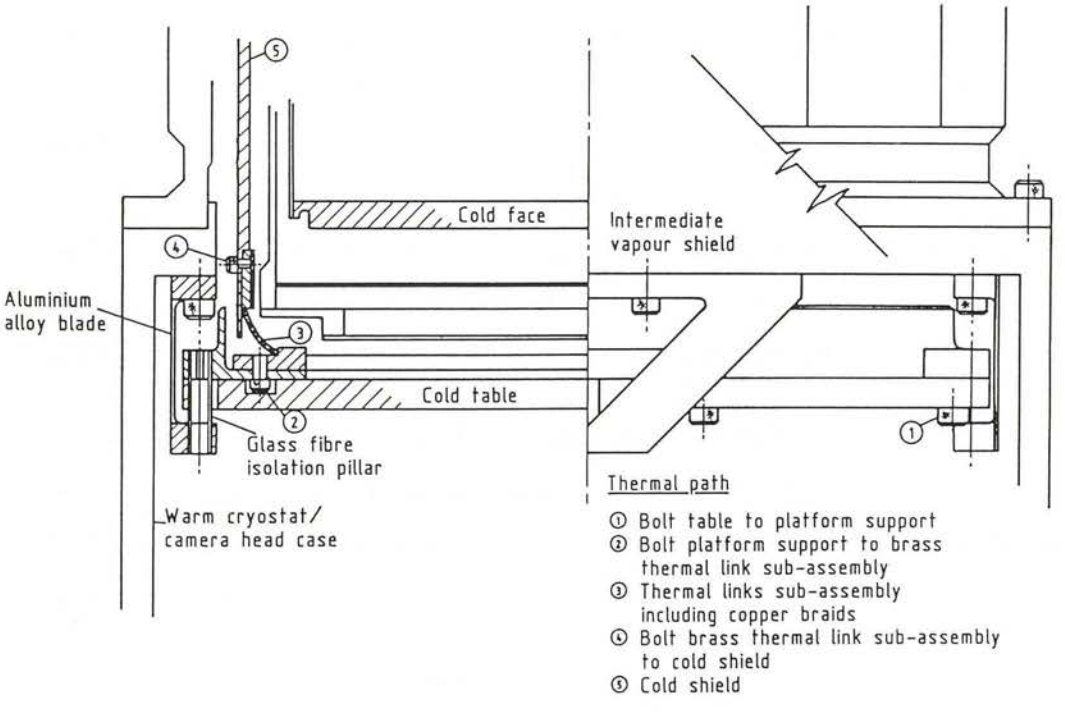


Figure 4.6 : Cross–section of the dewar/camera head interface.

4.3.3 Predicted Performance

Using an empirical estimate for the effective emissivity of the MLI, the model thermal discussed here has been used to predict the cryostat hold–times to within 20% of those measured, and to obtain a cold enclosure temperature below the specified upper limit of 85K (see Section 4.5). The compromises imposed on the thermal design by the requirement of low flexure and the need for ease of access and assembly (more bolted joints and accompanying contact resistances) are discussed in the following section on cryogenic mechanisms.

4.4 Cryogenic Mechanisms

4.4.1 Dewar/Camera Head Interface

The interface between the dewar and the camera head is illustrated in Figures 4.6 and 4.9. Even though mounting the cold table directly on the LN_2 cold shield cylinder would provide sufficient stiffness to satisfy the imaging requirements, this approach was not feasible because of the thermal contraction stresses involved. For this reason the cold table is mounted on the warm cryostat case by floatable supports which allow for contraction of the cold table as it cools from room temperature to about 80K. The cold table suspension system consists of 4 assemblies each comprising a fibre–glass thermal isolation pillar in series with 2 aluminium alloy spring blades (an "A" frame). Each of these assemblies is connected to the cold table as shown in Figure 4.6. Compliance of the blades in the direction normal to the axis of thin cross–section relieves thermal contraction stresses, while the thick cross–section provides the necessary stiffness against variable loading which results from attitude changes of the camera on the telescope.

Each blade is in effect a cantilever with one end fixed (to the warm cryostat case) and the other end free but guided (by attachment to the cold table). (For a derivation of the following cantilever equations see Roark 1956.) In this configuration the deflection δ_x due to a load W applied to the free end (see Figure 4.7) is given by :

$$\delta_x = \frac{WL^3}{12EI_z} \quad \text{where} \quad I_z = \frac{bd^3}{12} \tag{4.19}$$

and where L is the length of the blade

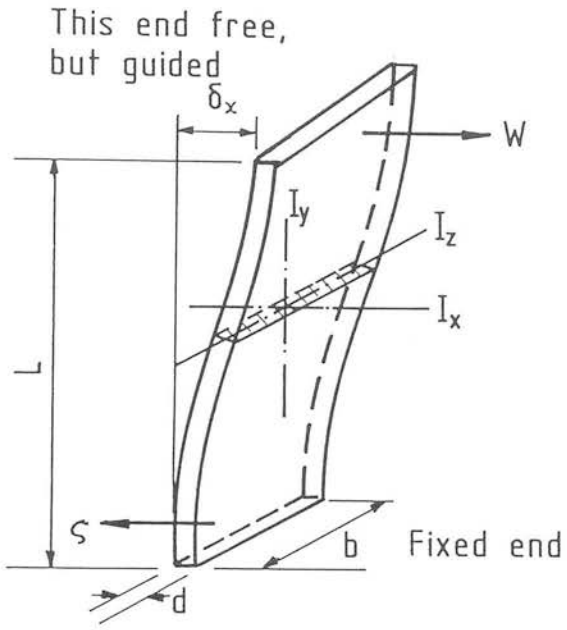


Figure 4.7 : Cantilever. Fixed at one end, load W is applied to the free but guided end. W produces a flexure δ_x at the free end and stress ς at the fixed end.

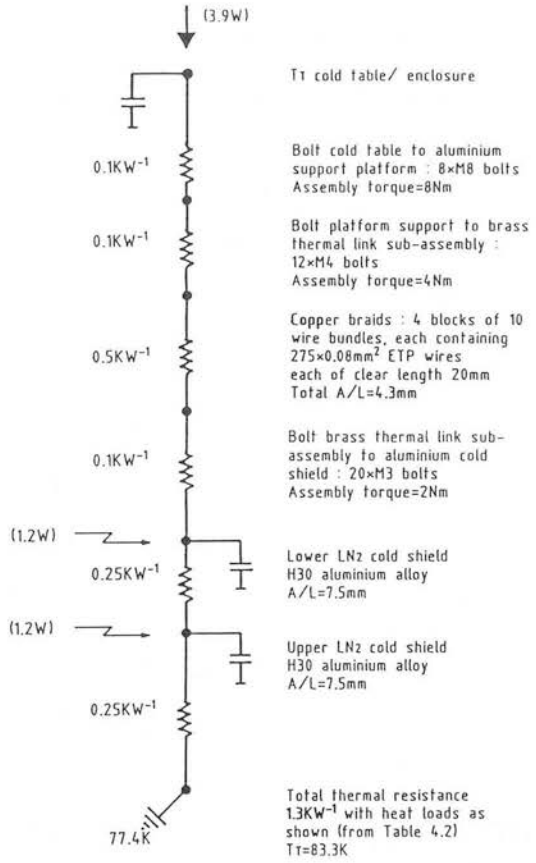


Figure 4.8 : Dewar/camera head interface thermal path (symbols as Figure 4.4).

E is Young's Modulus

I_z is the moment of inertia of the blade section (area bd) about the axis of bending (axis z in this case)

The (equal and opposite) stress at the joints is given by :

$$\varsigma = \frac{WLd}{4I_Z} = \frac{3\delta_x Ed}{L^2} \tag{4.20}$$

which must be safely less than the maximum yield stress of the material (165 Nmm^{-2} for aluminium alloy at 77K) for the joint to survive the applied load. On cooling to 80K the table contracts 1.2mm ($2\delta_x$) resulting in a stress at the A frame joints of $40\mathrm{Nmm}^{-2}$, well within that allowable. The cold optical enclosure and its components has a design weight of not more than 100N. If this load is applied to the A frames in the same direction as the thermal contraction, the flexure is about 2.5mm. (Such a change in loading would result if the camera was tipped through 90°.) Since the optical beam entering the cryostat is collimated, this movement would not cause an image shift on the detector array, but instead some loss of signal due to occultation of the the beam by the Lyot stop, which is rigidly fixed to the table. However, in the worst case scenario outlined here, the load is at the same time applied normal to the axis of thick cross–section of two of the other A frames such that (cf Equation 4.19) :

$$\delta_z = \frac{WL^3}{12EI_x} \quad \text{where} \quad I_x = \frac{db^3}{12} \tag{4.21}$$

In this direction the moment of inertia is considerably increased and the flexure correspondingly reduced to only about 1 μm (joint stress 4 Nmm^{-2}). For the fibre–glass rods :

$$\delta_x = \delta_z = \frac{WL^3}{12EI_Z} \quad \text{where} \quad I_x = I_z = \frac{\pi}{64}(\phi_o^4 - \phi_i^4) \tag{4.22}$$

and where ϕ_o is the outer diameter of the rod

ϕ_i is the inner diameter of the rod

From this equation it is apparent that low flexure results from short rods with large cross–sectional areas – exactly the opposite to the requirements of thermal isolation. Optimisation has resulted in a length L = 15mm and a cross–sectional area A = $15\mathrm{mm}^2$. The flexure of the fibre–glass pillars, due to a 100N load, is also very small (3 μm).

Good thermal coupling yet mechanical decoupling of the cold table with the LN_2 cold shield is brought about by using flexible copper braids. Access and assembly constraints have resulted in 3 bolted connections each adding more thermal contact resistance. Eight M8 bolts secure the cold table to the aluminium alloy platform support assembly, into which the fibre–glass pillars are glued. For each bolt, the area of contact between the cold table and the platform is conservatively assumed to be twice that covered by the diameter of the bolt head (beyond this warping tends to separate the surfaces). The joint is assembled by tightening the stainless steel bolts to a torque which is about half the yield torque. Using standard engineering tables of tightening torque against induced load, the joint pressure is estimated, and from Figure 4.2 an estimate of the joint conductance (or resistance) can be made. Allowance is also made for the reduction of induced load on cooling caused by the differential contractions of stainless steel and aluminium. The need for 8 bolts arises from the requirement to keep the joint resistance below $0.1KW^{-1}$.(The joint resistance is inversely proportional to the area of contact and therefore to the number of bolts.) Similarly this requirement has resulted in 12 M4 bolts and 20 M3 bolts at the other joint interfaces. Since each joint is assembled with a thin layer of silicone vacuum grease, a joint resistance of $0.1KW^{-1}$ is likely to be an upper limit. (Measurements of the cold enclosure temperature and estimates of the heat flowing through the dewar/camera head interface indicate that each resistance cannot be significantly greater than $0.1KW^{-1}$.)

In order to meet the demanded low thermal resistance, the standard 1.5mm thick aluminium alloy (H30) LN_2 cold shield wall of the MD1810 dewar ($A/L = 5.5$mm) was replaced with a 4mm thick wall ($A/L = 14$mm). Figure 4.8 illustrates the thermal path connecting the cold table to the LN_2 vessel. The total thermal resistance of the interface is calculated to be $1.3KW^{-1}$ and using projected estimates of the heat flowing through it (see Table 4.2), the temperature of the cold table is calculated to be about 83K.

4.4.2 Detector and Focus Assembly

Design of the detector assembly is governed by similar criteria to those applied to the cold enclosure, namely rigidity, low heat input,ease of assembly and access. An additional complication is the need for focusing. The detector assembly is mounted on

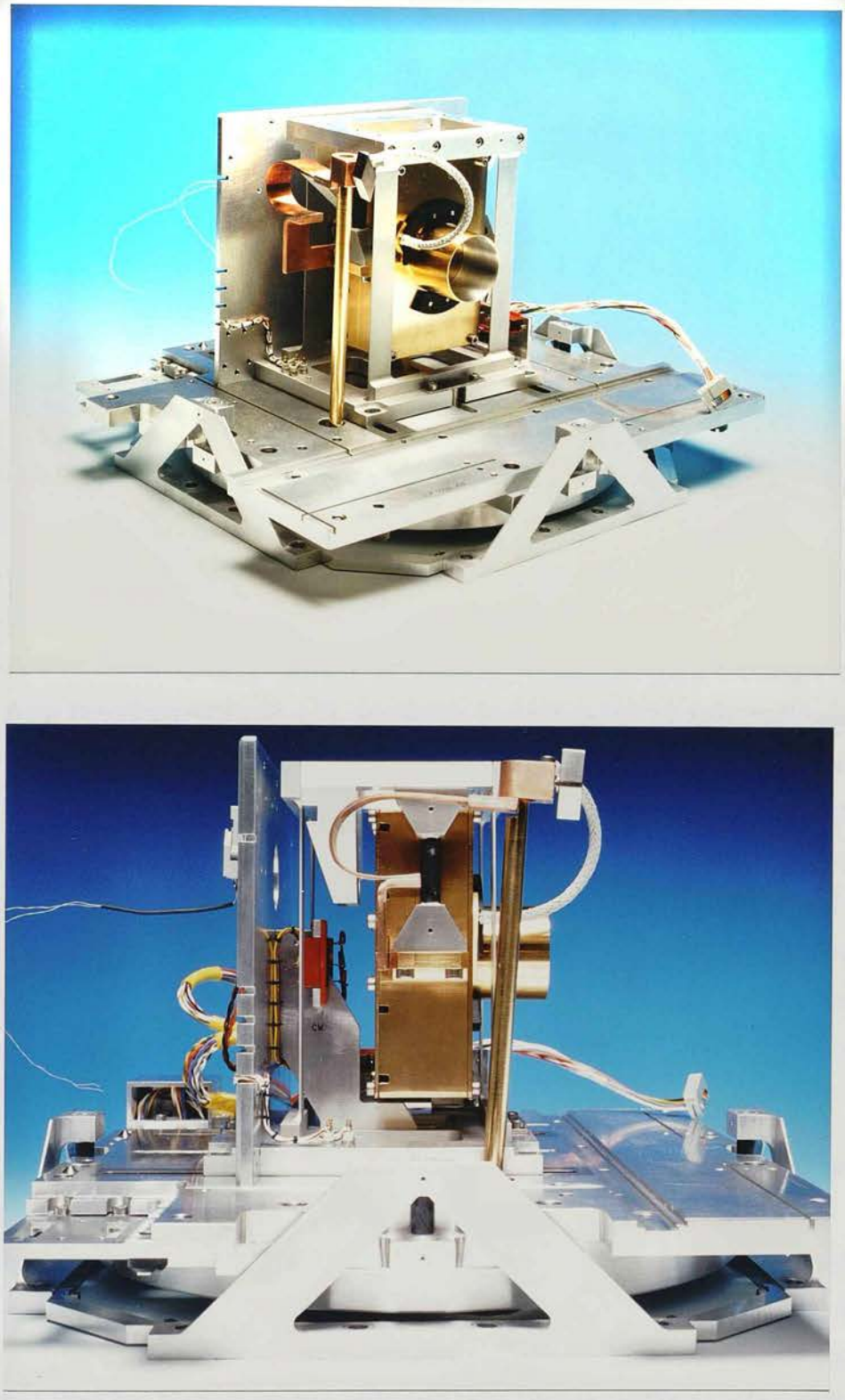


Figure 4.9 : Partial assembly of the detector housing, focus assembly, cold table, and cold table support structure. The gold–plated detector housing is rigidly supported from the focus assembly by 2 fibre–glass thermal isolation pillars. A screw mechanism (not shown) pushes on the 4 aluminium alloy spring blades to provide the focus action. This assembly is mounted on the cold table which is rigidly supported on the warm cryostat case by 4 A frames, each in series with a fibre–glass thermal isolation pillar. The detector housing is connected to the LHe cold face by a flexible copper foil and a rigid copper cold finger.

a four blade suspension system by two rigid fibre–glass thermal isolators, very similar to the cold table suspension system (see Figure 4.9). Focusing is accomplished by a screw mechanism which moves the detector ±3mm in the direction normal to the axis of thin cross–section of the blades. Rigidity in this direction is due to the the screw mechanism which is tied to a "thrust"plate. A fibre–glass focus shaft connects to the screw mechanism outside the cold baffles (see Figure 3.7), so that 1–5 μm photons from the warm regions of the shaft cannot be seen by the detector. Rigidity normal to the focus direction is provided by the blades' axis of thick cross–section, as was the case for the cold table suspension system.

In contrast to the cold table, flexure of the detector assembly relative to the optical beam does result in an image shift on the array. As has already been shown, flexure of the cold table relative to the optical beam is at most a few microns, and so we can consider flexure of the detector relative to the optical table as essentially equivalent. The detector assembly has a weight of 5N and if the lateral loading on it changes by 5N, then from Equations 4.21 and 4.22, the blades and pillars flex by 8 μm and 2 μm respectively. Such a change in loading could result when objects are tracked through the telescope zenith for 6 hours (90° attitude change), and with IRCAM mounted on either the east or west port of the telescope. The total shift of 10 μm is equal to $1/8^{\text{th}}$ of a pixel. Since the dark current alone limits single on chip integrations to less than $1\frac{1}{2}$ hours, the maximum flexure is about $1/20^{\text{th}}$ of a pixel, negligible compared to the affects of variable seeing and telescope tracking errors. Indeed IRCAM is currently mounted on the north port of the telescope, such that any change in loading is along the focus direction, where the mounting is even more rigid.

The detector array is mounted on a copper heat sink which is connected to the cold finger by flexible high thermal conductivity copper foils (see Figure 4.9). This arrangement allows for fine (filter change) and gross (lens change) movements of the detector. As illustrated in Figure 4.10 a clamp plate presses the detector chip carrier into electrical contact with the detector printed circuit board (PCB), and indium foil ensures a very low thermal contact resistance with the copper block at low clamping pressures. This assembly is enclosed in, and secured to, an aluminium box–like housing, and the complete structure forms the Detector Assembly (or Cold Stage).

Differential thermal contraction relative to the cold enclosure optical axis is minimal

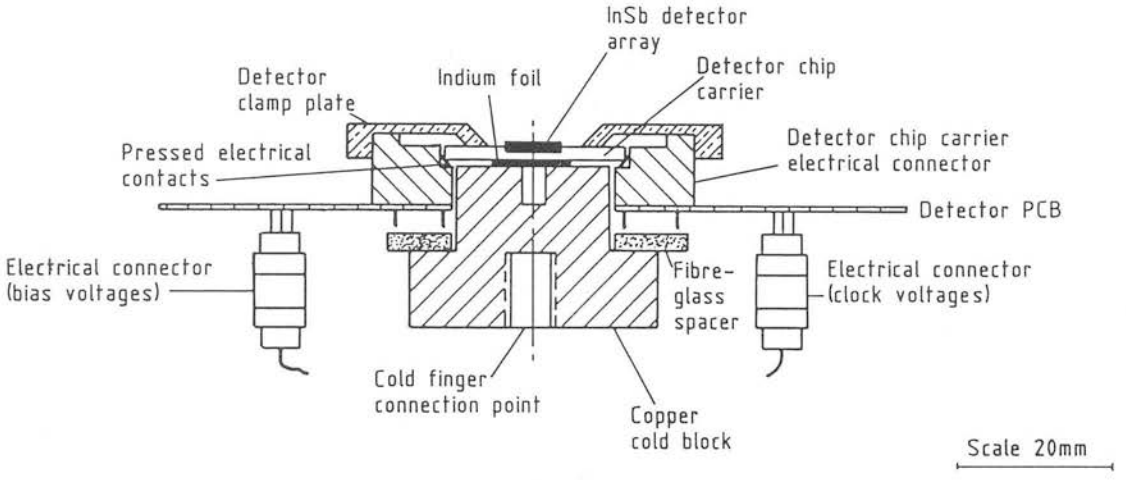


Figure 4.10 : Detector array mount (SBRC design). The SBRC InSb array is mounted in a 68 pin leadless chip carrier, which is pressed into thermal contact with a copper cold sink, and into electrical contact with the detector PCB, by a clamp plate.

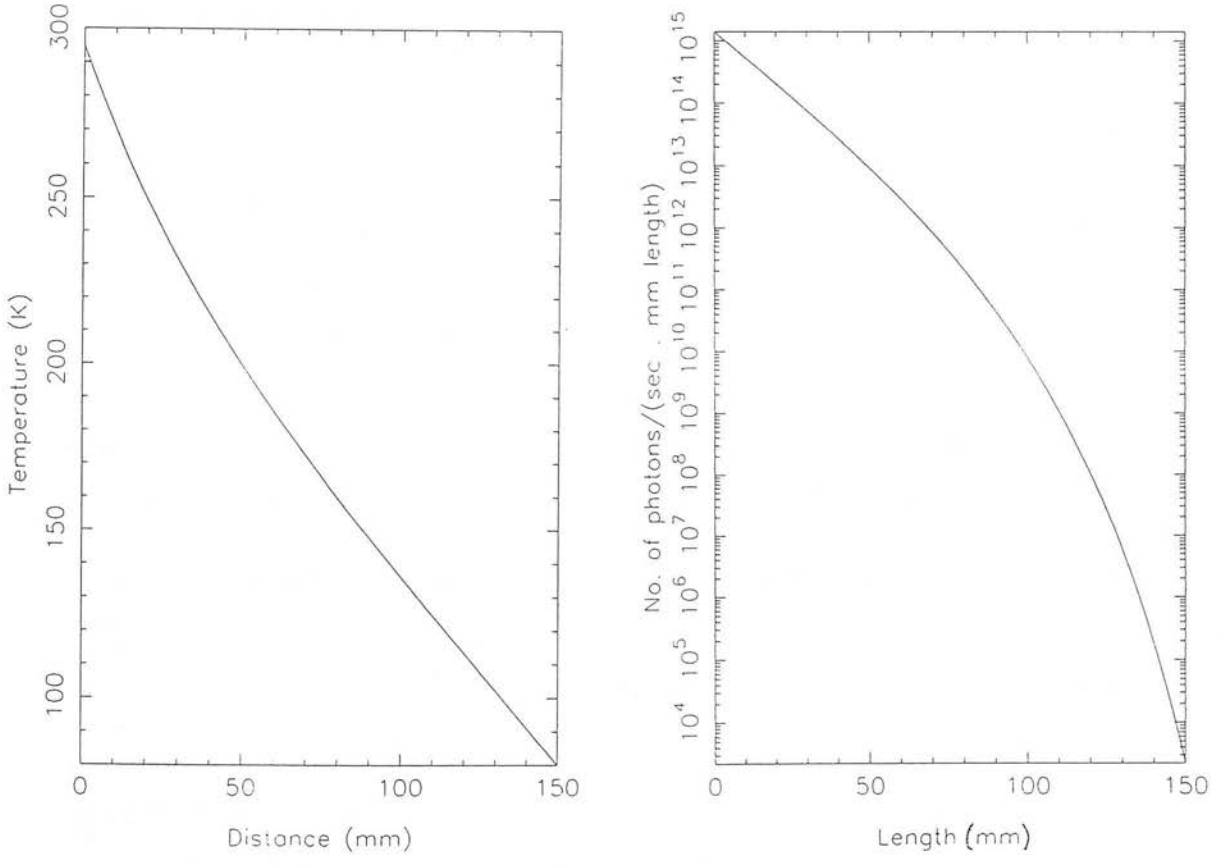


Figure 4.11 : Temperature gradient in the filter wheel fibre–glass drive shaft, and the number of 1–5.1 μm photons emitted by the shaft. The shaft is assumed to have an emissivity of $\varepsilon = 0.5$, a length of $L = 150$mm, a cross–sectional area of $A = 15\text{mm}^2$, and a diameter of 7.0mm.

since all the optical components (Lyot stop, filter wheels and lenses), are mounted on the cold table by aluminium alloy structures. Although the detector assembly is also suspended on aluminium alloy blades, the fibre–glass support pillars introduce an unavoidable 75 μm contraction differential. Heat inputs to the detector assembly come from four main sources, conduction through wiring and the support pillars, thermal radiation from the cold enclosure, and detector power dissipation. About 30 wires need to connect with the array. To reduce electrical noise each wire is screened with an earth wire, making about 60 wires in all. Also, the array must be capable of being read out at frequencies of ~100kHz (3596 pixels in ~50ms), and for good performance this requires short wires. However, from a thermal standpoint the wires should be of small cross–section and long. Consequently, wiring is a compromise between electrical and thermal performance. As was the case with the cold table fibre–glass supports, the A/L ratio of the detector assembly fibre–glass supports is a trade–off between flexure and heat load.

4.4.3 Drive Shafts and Filter Wheels

Because of the isolating properties of bearings, the aluminium alloy filter wheels are thermally connected to the cold table by high conductivity flexible copper foils (see Figure 3.6). The foils are wound much like clock springs and allow the wheels to freely rotate through $1\frac{1}{2}$ revolutions, although mechanical datum stops restrict movement to precisely one revolution. The wheels are turned by fibre–glass drive shafts connected to stepper motors mounted on the outside of the cryostat. To allow for differential thermal contraction, the drive shafts connect to the stepper motor vacuum feedthroughs via Oldham couplings which accommodate the 1mm of relative movement.

The angular deflection θ of a rod subject to a torsional load W_T is given (Roark 1959) by :

$$\theta = \frac{W_T L}{G I_y} \quad \text{where} \quad I_y = \frac{\pi}{32}(\phi_o^4 - \phi_i^4) \tag{4.23}$$

and where L is the length of the rod

G is the shear modulus

I_y is the moment of inertia about the axis of the rod

ϕ_o is the outer diameter of the rod

ϕ_i is the inner diameter of the rod

The fibre–glass shafts which drive the filter wheels have outer and inner diameters of 7mm and 5.5mm respectively (A=15mm^2), and are 150mm long. For the same torsional rigidity stainless steel shafts of the same length and outer diameter require an inner diameter of 6.8mm. On the basis of the solution to Equation 4.5, Figure 4.11 shows the temperature gradient in the fibre–glass shafts, and the number of 1–5 μm photons emitted from the shaft due to the temperature gradient. For an emissivity of 0.5 the radiated and conducted heat loads are 70mW and 2mW respectively, which compares to 120mW and 15mW for a torsionally equivalent stainless steel shaft. Since stainless steel shafts of this thickness (0.2mm wall) are easily damaged, fibre–glass shafts were chosen. The detector is shielded from the large numbers of infrared photons emitted from the shafts by the cold baffles.

4.5 Cryogenic Performance

The estimated magnitudes of the heat inputs to the LN_2 and LHe vessels of the cryostat are listed in Tables 4.2 and 4.3, and on this basis the hold–times are calculated to be 25 hours and 44 hours respectively. With the heat loads as shown the best estimate of the cold table temperature is 83K (see Figure 4.8). The cryogenic specification called for both the cold enclosure and detector assembly to take no longer than 12 hours to reach 100K. Pre–cooling of the detector assembly is done by placing LN_2 in the inner vessel. Once the detector assembly has cooled to below 100K, cooling is continued by replacing the LN_2 in the inner vessel with LHe. For the most efficient use of LHe it is better to allow the detector assembly temperature approach 80K before continuing. Measured cooling times of various components within the camera are plotted in Figures 4.12 and 4.13. The cold table and detector assembly (cold stage) take about 10 hours to reach 100K and the detector assembly requires a further 3–4 hours to reach a stable operating temperature of 35.0$\pm$0.1K. It is relatively easy to warm–up the camera ($\sim$12 hours), make an adjustment (eg change a lens), attach to a pump and evacuate the vacuum space ($\sim$6 hours), and then cool to the operating temperature ($\sim$18 hours), well within two days. Fast cycling times such as this are highly desirable, particularly during the instrument test phase when unforeseen problems are more likely to occur. During testing in the laboratory LN_2 and LHe hold–times of 30 hours and

48 hours were routinely obtained. Since the corresponding heat loads were slightly lower than predicted and the cold table temperature was found to be 83K, the thermal resistance of the dewar/camera head interface turned out to be higher the expected, but still within specification.

In the telescope dome the ambient temperature (~273K) and air pressure (0.6 atmospheres) are lower than in the sea level laboratory. Consequently the cryogenic performance of the camera on the telescope is in two respects better than in the laboratory. Due to the lower ambient temperature the heat load radiated onto the LN_2 cold shield is about 25% less and the boiling point of the LN_2 in the outer vessel falls to 74K. Both these effects result in a lower cold table temperature. The measured performance is consistent with this, although when the camera is on the telescope the outer vessel hold–time is affected by tipping the LN_2 up the vent tube. From high spectral resolution observations using the Fabry Pérot étalon, there is no evidence to suggest the presence of any significant thermal background within the camera.

Table 4.2: **Heat Load on the LN_2 vessel** : capacity=4.5 litres (201Whr)

Source	*HeatLoad/W*	*Comments*
Dewar	3.5	Measured, mostly radiation
Camera head radiation	3.4	A_2=0.42m^2,A_1=0.33m^2 F_{2-1} = 0.025, T_2=290K,T_2=80K
Table suspension	0.5	Fibre–glass A/L=4mm
Detector pre–amp	0.2	Pre–amp at 120K
Drive shafts:		
filter wheels (2)	0.15	Fibre–glass A/L=0.2mm
focus (1)	0.07	Fibre–glass A/L=0.3mm
Wiring	0.1	156 x 0.012mm^2 (SWG 40) Constantan, each 100mm long
Gas switch capillary	0.05	Stainless steel A/L=2x10^{-2}mm
Window radiation	<0.05	ε=0.1, diameter=40mm
Gas conduction (camera head)	0.005	Vacuum pressure=10^{-5}mbar
TOTAL	**8.0**	**Hold–time = 25 hrs**

Table 4.3: **Heat Load on the LHe vessel** : capacity=4.0 litres (2960mWhr)

Source	*HeatLoad/mW*	*Comments*
Dewar	15	Measured
Detector suspension	25	Fibre–glass A/L=1.2mm 85K to 35K
Detector power dissipation	10	Detector at 35K
Wiring pre–amp (120K to 35K) digital (85K to 35K) analogue (85K to 35K) gas switch (85K to 4K)	5	Constantan, 0.012mm^2 (SWG 40) 2x2 wires each 200mm long 2x19 wires each 350mm long 2x11 wires each 350mm long 2x6 wires each 200mm long
Radiation onto detector assembly	3	A_2=0.1m^2, A_1=0.03m^2 F_{2-1}=0.03, T_2=85K, T_1=35K
Z pump heater	7	Maintains detector at 35K
Gas switch capillary	2	Stainless steel A/L=5x10^{-3}mm
TOTAL	**67**	**Hold–time = 44 hrs**

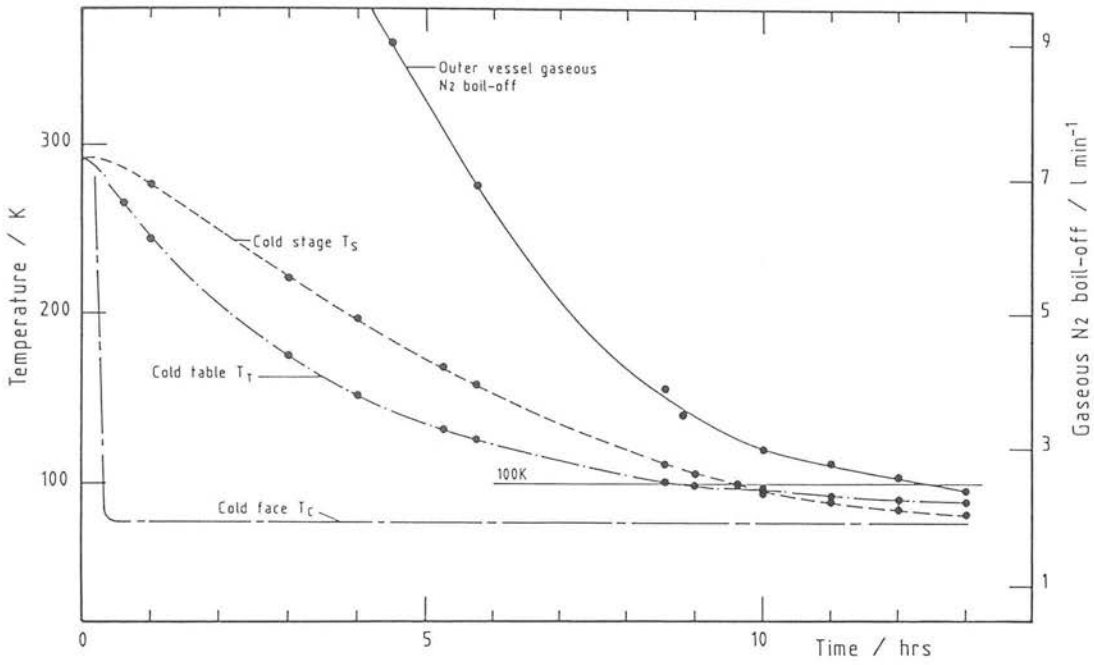


Figure 4.12 : Cooling with LN_2.

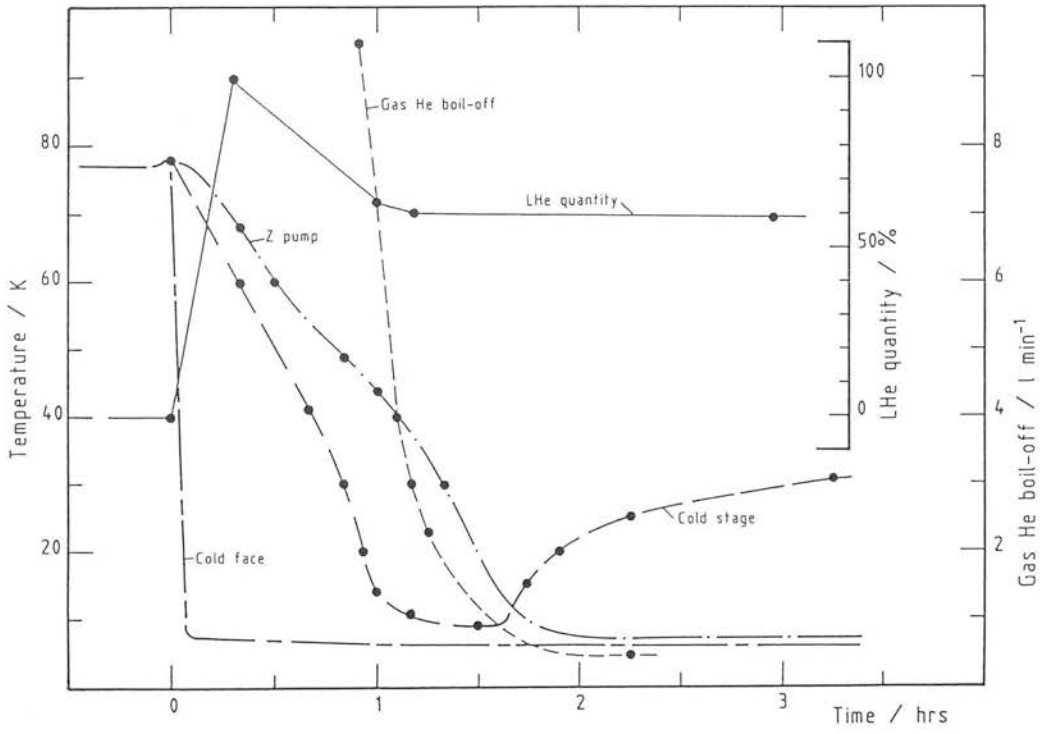


Figure 4.13 : Cooling with LHe 24 hours after initial cooling with LN_2.

Chapter 5

Detector Temperature Control

5.1 Introduction

Design of the detector temperature control system was driven by the following criteria :

- A liquid helium (4 litre capacity = 2960mW Hr) hold-time of at least 36 hours (equivalent to a 80mW heat leak) must be maintained while the detector is at its operating temperature.

- The detector must be able to operate at any static temperature in the range 10–60K and be stable to ± 0.1K.

- The supply of LHe cannot always be guaranteed at the summit of Mauna Kea. However, since broad band imaging can still be done when the detector is at solid nitrogen temperatures ($\sim$ 47K), albeit at reduced S/N (see Chapter 2.3), it is highly desirable for the cold stage to operate as close as possible to the cold–face temperature when there is SN_2 in the inner vessel.

- Cool-down of the 0.5Kg cold stage to a nominal operating temperature of 30K should take no longer than about one day. On the other hand, if cooled too rapidly detectors may fail due to thermal shock. Indeed the effects of accumulated temperature cycling can also cause detectors to fail. The SBRC 62x58 InSb array + DRO has a recommended cooling/warming rate of 1K min^{-1} in the region 300–77K, and so the minimum cooling time is about 4 hours. Much higher rates

are allowable below 77K because thermal contraction is greatly reduced at these temperatures.

Two schemes were considered. The first involved a simple thermal potentiometer in which the cold stage was connected to the liquid helium cold-face by a thermal (wire) link (Rayner 1984). Detector temperature control was achieved using an electrical (resistor) heater mounted on the cold stage. However, to keep the liquid helium heat leak below the specified 80mW, more than one fixed thermal link was required to cover the entire temperature range and the link would have proved very difficult to change remotely, that is, without warming up the cryostat. Also, the cool–down time and the solid nitrogen operating mode demanded some form of removable thermal–short to the cold–face. Given the potential difficulties of a mechanically actuated thermal short, a passive gas switch was favoured. A prototype VCGS was built and the variable conductance mode demonstrated.

The physical principle on which the operation of the gas switch is based is the conduction of heat by a gas. The thermal conductance of the gas switch can be controlled by varying the pressure of a gas contained within it. In the following section we will show how the conduction of heat through a gas is related to its pressure.

5.2 The Conduction of Heat by a Gas

The conduction of heat by a gas depends not only on the nature of the gas but also on its temperature and pressure. Two distinct regions must be considered (see for example, Roberts and Miller 1960) :

1. $d \ll L$

2. $d \gg L$

where L is the distance separating the temperature boundaries, and the mean free path (mfp)in the gas d is conveniently obtained from tabulated values of viscosity using the formula :

$$d = 1.14\text{x}10^{-4} \frac{\eta}{P} \left(\frac{T}{M} \right)^{\frac{1}{2}} \tag{5.1}$$

Table 5.1: **Mean Free Path** (d in mm, at $P = 10^{-3}$ mbar)

$Temperature(K)$	$Helium$	Air
4	1.4	–
20	8.7	–
80	42	11
300	195	66

where d is in mm

T is the absolute temperature of the gas in K

η is the viscosity of the gas in micropoise (10^{-6}g cm^{-1} s^{-1})

M is the molecular weight

Table 5.1 lists selected mean free paths for helium and air calculated from Equation 5.1

5.2.1 Mean free path $\ll$ separation of temperature boundaries

Consider a gas contained within two infinite parallel plates at temperatures T_1 and T_2 and that a steady state exists in which heat is conducted across the gap at a constant rate H and there is no convection. In Figure 5.1 a parallel plane is drawn at a distance x from the plane T_2. Since the number of *hot* gas molecules crossing the plane from above is equal to the number of *cold* gas molecules crossing the plane from below, there is a net transfer of energy but no net transfer of gas. The number of gas molecules which in one second cross unit area of this plane in a direction making an angle θ is from the geometry $\frac{1}{2}n\bar{c}\cos\theta\sin\theta d\theta$ where $\bar{c}$ is the mean velocity of the gas molecules and n is the number of molecules per unit volume. A gas molecule crossing this plane at an angle θ will have travelled on average a distance d equal to the mean free path and its energy will be :

$$E - d\cos\theta\frac{\partial E}{\partial x}$$

Thus the total energy flow across the plane per unit area per time :

$$\frac{H}{A} = \int_0^{\pi}\left(E - d\cos\theta\frac{\partial E}{\partial x}\right)\frac{1}{2}n\bar{c}\cos\theta\sin\theta d\theta = -\frac{n\bar{c}d}{3}\frac{\partial E}{\partial T}\frac{\partial T}{\partial x}$$

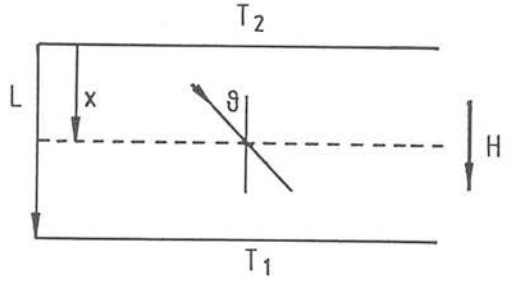


Figure 5.1 : Heat transfer by gas conduction between two parallel infinite plates.

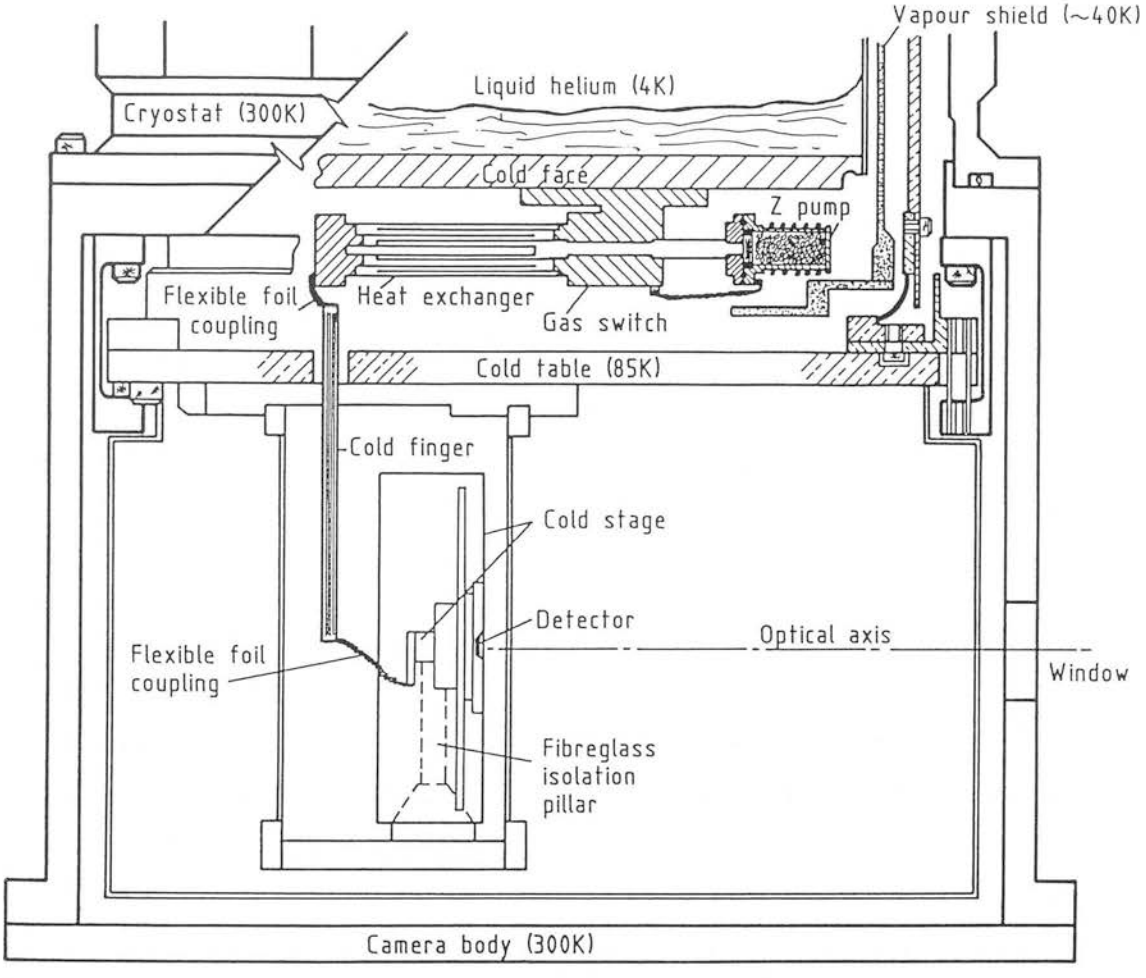


Figure 5.2 : Configuration of the detector array cryogenic temperature control system.

Now :

$$\frac{\partial E}{\partial T} = mC_v \quad \text{and} \quad \frac{H}{A} = -k\frac{\partial T}{\partial x}$$

where m is the mass of a gas molecule

C_v is the specific heat at constant volume

k is the thermal conductivity of the gas

Therefore :

$$H = \frac{A}{L}\int_{T_1}^{T_2} k dT \quad \text{where} \quad k = \frac{1}{3}ndcmC_v \tag{5.2}$$

Mean free path d is inversely proportional to pressure but molecular number density n is proportional to pressure, and so the product nd in Equation 5.2 is pressure independent. Therefore when the mfp is much less than the boundary separation the gaseous thermal conductivity does not depend on pressure, that is, heat flow is independent of gas pressure.

5.2.2 Mean free path $\gg$ separation of temperature boundaries

In this case any gas molecule leaving plate T_2 will reach plate T_1 without on average hitting any other gas molecules and vice versa. A molecule leaving the upper plate impinges on the lower plate with an energy corresponding to T_2, and will leave the lower plate not at an energy corresponding to T_1, but with an energy corresponding to some temperature T_1' intermediate between T_2 and T_1. The degree to which the molecule has come into thermal equilibrium with the surface is known as the Accommodation Coefficient, a, where :

$$a = \frac{T_2 - T_1'}{T_2 - T_1}$$

For most metal surfaces the accommodation coefficient varies between 0.2 and 0.8. The value depends on the physical state of the surface and the nature of the gas. For a rough surface a can approach unity while for a very clean surface, such as tungsten wire after flashing, values of $a < 0.1$ are obtained (White 1979). However, for surfaces and gases usually encountered in cryostats, $a \sim 0.5$ is a reasonable approximation. From simple kinetic theory the number of gas molecules impinging on unit area of the lower plate is $\frac{1}{4}n_2\bar{c}_2$ and each of these carries energy mC_vT_2. Similarly for the upper plate. Therefore the net heat flow to the lower plate is :

$$\frac{H}{A} = \frac{1}{4}n_2\bar{c}_2 mC_v T_2 - \frac{1}{4}n_1\bar{c}_1 mC_v T_1$$

Provided no condensation occurs, the numbers of molecules reaching and leaving each plate must be the same, thus :

$$\frac{H}{A} = \frac{1}{4} n \bar{c} m C_v (T_2 - T_1)$$

In this case it is meaningless to define a thermal conductivity for the gas since the rate of heat transfer is independent of the temperature gradient, although dependent on the temperature difference. Since the molecular number density n is directly proportional to the pressure, the heat flowing through the gas is also proportional to pressure.

A more rigorous treatment (Corruccini 1959) of the heat flow between concentric spheres, coaxial cylinders, or parallel planes, gives :

$$\frac{H}{A_1} = a \frac{\gamma + 1}{\gamma - 1} \left(\frac{R}{8\pi} \right)^{\frac{1}{2}} \frac{P}{(MT)^{\frac{1}{2}}} (T_2 - T_1) \quad \text{where} \quad a_0 = \frac{a_1 a_2}{a_2 + a_1 (1 - a_2) A_1 / A_2} \tag{5.3}$$

where a_0 is the overall accommodation coefficient and A_1/A_2 is just a geometrical configuration factor (see Chapter 4.2.1 and Equation 4.20)

H is net heat flow per unit time

γ is C_p/C_v, the specific heat ratio of the gas, assumed constant

a_1 is the accommodation coefficient of the inner surface

a_2 is the accommodation coefficient of the outer surface

R is the molar gas constant

P is the gas pressure

M is the molecular weight of the gas

T is the effective gas temperature

A_1 is area of inner surface, temperature T_1

A_2 is area of outer surface, temperature T_2

The constant $(R/8\pi)^{\frac{1}{2}}$ has the value 0.2426 for H in W cm^{-2}, P in mm Hg,and T in K.

Corruccini (1959) points out that T in the denominator of Equation 5.3 is an effective temperature of the non-equilibrium gas which has a value intermediate between T_1 and T_2 but is not the average. However, since in our case surface areas A_1 and A_2 are equal, the effective gas temperature is approximately equal to the average (Corruccini 1959). When $d \gg L$, the thermal transpiration effect operates and the local pressure varies with local temperature according to the law, $PT^{\frac{1}{2}} =$ constant. When P is measured at an external pressure gauge which is connected to the gas switch by a capillary

tube of bore much less than the mfp, the temperature at the gauge should be inserted for T, even when this temperature lies outside the limits T_1 and T_2. However, when the mfp within the gas switch is being considered the local pressure and effective gas temperature must be used.

It is apparent from Equation 5.3 that the ambient gas pressure is inversely proportional to the temperature difference (T_2-T_1). So, as long as the unconfined mfp of the gas is greater than the gas gap, the temperature drop across the gas switch may be changed simply by varying the pressure. This is analogous to changing the voltage drop across a resistor by varying resistance rather than current.

Helium is the only element which will remain gaseous over the entire operating range of 10K and above. Its use as an exchange gas to make and break thermal contact in cryogenic apparatus is well documented (see for example, Nast *et al.* 1982 and, Hilberath and Vowinkel 1983). The use of cryogenic sorption pumps as exchange gas reservoirs is also well known, but Torre and Chanin (1984) were the first to suggest operating a cryopump at temperatures intermediate to the normal low (gas completely adsorbed) and high (gas desorbed) temperatures. In this manner the exchange gas pressure is a function of cryopump temperature and the gas switch operates as a variable conductance.

Equation 5.3 may be re-written explicitly for helium gas:

$$H = 3.64 a_0 P (T_2 - T_1) A / T^{\frac{1}{2}} \tag{5.4}$$

where H is in mW

P is in mbar

A is in mm^2 ($A = A_1 = A_2$)

Of course the heat flow H must be minimised to conserve liquid helium. The temperature drop ($T_2 - T_1$) depends upon the chosen detector temperature, T_2 (ignoring for the moment the small gradient in the cold–finger which is required to connect the cold stage to the gas switch) and the cold–face temperature, T_1 (5K).Therefore, it follows that for a given heat flow and temperature drop, the exchange gas pressure, P, is inversely proportional to the surface area A. Consequently, relatively large areas require low exchange gas pressures in the variable conductance regime and vice versa. The implications of this are discussed further in the following sections.

Since the gas switch is to be used in both conductance regimes, we need to know where the transition occurs, that is, the relationship between the unconfined mfp in the helium gas and the gas switch parameters. Rewriting Equation 5.1 explicitly for helium gas :

$$d = 5.72\text{x}10^{-5}\eta T^{\frac{1}{2}}/P \tag{5.5}$$

Therefore combining Equations 5.4 and 5.5 gives :

$$d = 20.8\text{x}10^{-5}\eta a(T_2 - T_1)A/H \quad \text{when} \quad d \gg L \tag{5.6}$$

This relationship is used to design the gas switch heat exchanger.

5.3 Design of the Gas Switch

In the IRCAM instrument the cold stage is supported from the (85K) optical table by two low conductivity fibre glass tubes (see Figure 5.2). It is thermally connected by a high conductivity copper cold–finger to the 'hot' end of a cylindrical chamber, called the Heat Exchanger, which is made of stainless steel (a poor conductor). The 'cold' end of this chamber is mounted on the liquid helium cold-face (see Figure 5.3). The heat exchanger contains 4 concentrically separated copper cylinders, 2 hot and 2 cold (see Figure 5.4). Connected to the exchanger by a stainless steel tube is a smaller copper cell which contains the adsorbent material, zeolite (Z) pellets. This cryopump or "Z pump" is thermally linked to the cold–face by a copper wire. The VCGS is filled at room temperature to one atmosphere with helium gas through a heat sunk stainless steel capillary and the system closed by an external valve.

5.3.1 Heat Exchanger

The cold stage has a large thermal mass (enthalpy 80,000J at 300K) and so to speed cool–down to its operating temperature the heat exchanger must provide a thermal short-circuit to the cold–face. Maximum conductance occurs when the mfp (d) is much less than the gas gap (L). In this 'on' condition exchanger conductance is from Equation 5.2 proportional to A/L. Consequently, for fast cooling A/L must be maximised and the mfp kept much less than the gas gap. Cooling is done in two stages, first with liquid nitrogen in the inner vessel and then with liquid helium. This avoids wasting

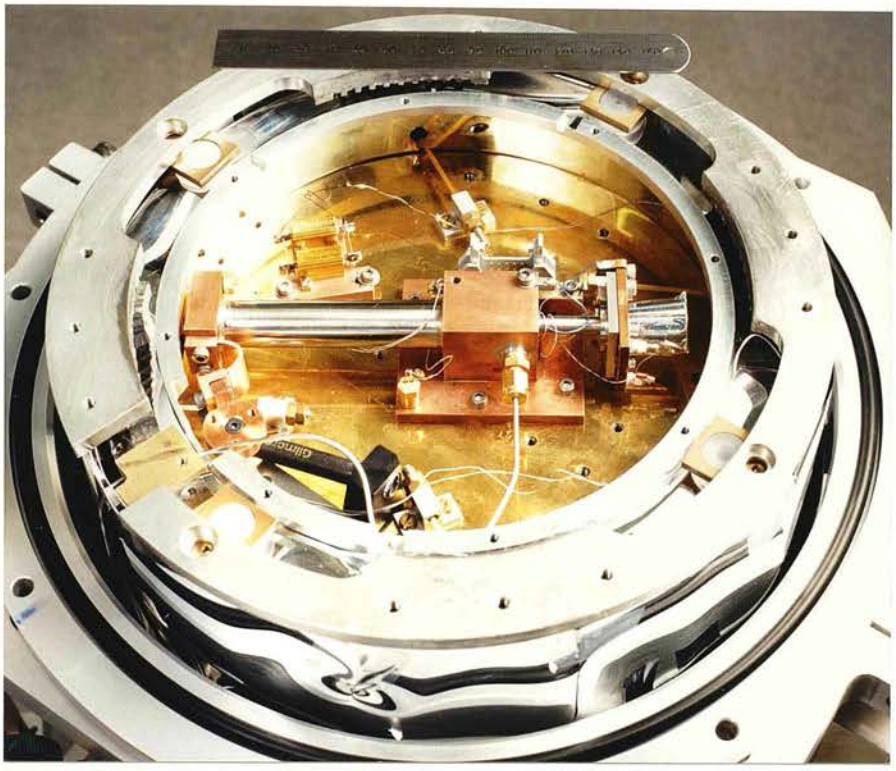


Figure 5.3 : Photograph of the Variable Conductance Gas Switch (VCGS). The gas switch is mounted on the LHe cold face, and a gas capillary tube connects it to a pressure gauge and fill/isolation valve mounted on the outside of the cryostat. The base of the cold finger is connected to the 'hot' end of the heat exchanger (see Figure 5.4) by a flexible copper foil which is rigidly supported on the cold face by a fibre–glass pillar.

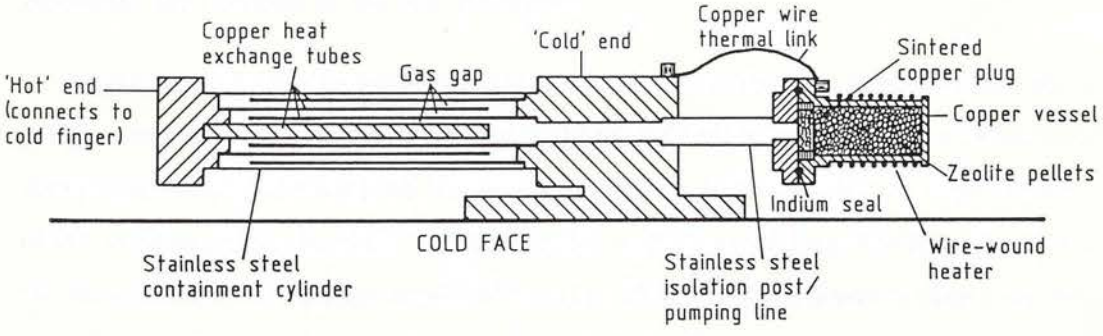


Figure 5.4 : The VCGS.

expensive liquid helium. During the first stage the Z pump never gets cold enough (it is always >77K) to significantly cryopump the helium exchange gas, while during the second stage of cooling the cold stage cools down before the Z pump is cold enough to pump significantly (<30K). Hence, during cooling the mfp is always much less than the gas gap, (see Figure 5.6). The exchange tubes themselves must also be good conductors and this is ensured by using thick copper tubes. In practice cooling is limited by the long length of the cold–finger, which is necessitated by the practicalities of detector positioning and accessibility.

Radiative heat flow is always less than conduction during cool–down. Heat flow due to radiative coupling of the exchange tubes is given by (see Chapter 4.2.1):

$$H_{rad} = \sigma \varepsilon A(T_2^4 - T_1^4) \tag{5.7}$$

where σ is the Stefan-Boltzmann constant

ε is the effective emissivity

It follows from Equations 5.2 and 5.7 that the ratio of radiative to conductive coupling is :

$$\frac{H_{rad}}{H_{cond}} = \frac{\sigma \varepsilon L}{k} \left(\frac{T_2^4 - T_1^4}{T_2 - T_1} \right) \tag{5.8}$$

In our case the maximum of this ratio occurs when T_2=300K, T_1=77K and ε=1. For helium gas at one atmosphere, the thermal conductivity k ~0.1 mW mm^{-1} K^{-1}, and from gas pressure control and space considerations (discussed further below), the gas gap L <1mm, resulting in H_{rad}/H_{cond} <0.02. Hence we can ignore radiative cooling. In fact we chose to design the exchange tubes for minimum emissivity in order to achieve minimum conductivity in the 'off' condition.

In the variable conductance regime governed by Equation 5.4, the exchanger must satisfy some very different requirements. When the switch is off, and the helium exchange gas completely adsorbed, the only significant heat flow path is through the wall of the stainless steel chamber. By making this wall thin and long, the heat flow from the detector and its thermal support structure (25mW), is sufficient to keep the hot end of the exchanger above 60K when the switch is off.

To maximise the on conductance, we could choose to make A large. However, as pointed out previously, the gas pressure is inversely proportional to A when operating in

the variable conductance mode. As A becomes larger the required pressure gets smaller and potentially more difficult to control. Also, A is limited by the space available; (cylinders are easy to make and are space efficient). An ideal solution would be to make the gas gap very small and make A as large as is consistent with the space and gas pressure control requirements. Unfortunately, the smaller the gas gap the harder the heat exchanger is to build.

A constraint on A/L imposed by the variable conductance mode can be obtained from Equation 5.6. In this mode the low temperature limit occurs when the unconfined mfp approaches the dimensions of the gas gap, that is, when the heat exchanger is moving from its intermediate conductance state ($d \gg L$), into its high conductance or on state ($d \ll L$). Re-arranging Equation 5.6 :

$$\frac{A}{d} = \frac{H}{20.8\text{x}10^{-5}\eta a_0(T_2 - T_1)} \text{ mm } d \gg L \tag{5.9}$$

In the limit where the mfp (d) approaches the gas gap (L) we can use this relationship to estimate a minimum value for A/L. For IRCAM we require a detector low temperature limit of about T_2=10K, at which temperature the heat flow from the detector and its thermal support structure, H=40mW. (When the temperature gradients in the cold finger and heat exchange tubes are small, $T_S \simeq T_H \simeq T_2$, $T_2 \simeq T_C$, and most of the temperature drop occurs across the gas gap ($T_2 - T_1$), see Figure 5.5) The cold–face temperature T_C=5K, and the effective gas temperature T=8K, at which temperature the viscosity η=18 micropoise and the overall accommodation coefficient $a_0 = 0.8$ (interpolated from Corruccini (1959)). Therefore replacing d by L in Equation 5.9, we obtain A/L >2700 mm.

We chose A/L=4000mm for the IRCAM gas switch, which from Equation 5.9, gives T_2=8.5K. A gas gap of L=1mm was decided upon primarily for ease of construction. The common surface area of A=4000mm^2 is achieved by 2 hot and 2 cold co-axial copper cylinders. These values are consistent with the cooling time (A/L) and Z pump exchange gas pressure control (A) requirements. Using Equations 5.4 and 5.5, Table 5.2 gives the model parameters for the heat exchanger performance in the variable conductance mode. Note that $d \gg L$ until 10K is approached.

Using Equation 5.2 ($d \ll L$) and allowing for the small thermal resistance of the copper exchange tubes, the on conductance of the heat exchanger when there is liquid helium in the cryostat inner vessel is about 25mW K^{-1}, resulting in a hot end temper-

Table 5.2: **Variable conductance mode model for A/L=4000mm**

T_2	T_1	T	$\eta(T)$	$a(T)$	H	H_s	H_g	P	P_o	d
60	5	33	48	0.3	25	25	< 1	$< 2\text{x}10^{-5}$	$3.8\text{x}10^{-4}$	>780
40	5	23	39	0.4	31	17	14	$3.3\text{x}10^{-4}$	$3.4\text{x}10^{-4}$	33
30	5	18	32	0.5	34	6	28	$6.5\text{x}10^{-4}$	$5.6\text{x}10^{-4}$	12
20	5	13	26	0.6	37	2	35	$9.6\text{x}10^{-4}$	$7.2\text{x}10^{-4}$	6
15	5	10	21	0.7	39	1	38	$1.2\text{x}10^{-3}$	$9.4\text{x}10^{-4}$	3
10	5	8	18	0.8	40	0	40	$2.0\text{x}10^{-3}$	$1.8\text{x}10^{-3}$	1.5
6.5	5	6	-	-	40	0	40	-	-	$\ll$L

where A is the common surface area of the exchange tubes (4000mm^2)

L is the gas gap (1.0mm)

T_2 is the temperature of the hot end of the heat exchanger (K)

T_1 is the temperature of the LHe cold–face (K)

T is the effective gas temperature (K)

$\eta(T)$ is the viscosity of the exchange gas in micropoise

$a(T)$ is the accommodation coefficient

H is the total heat flowing through the exchanger (mW)

H_s is the heat flowing through the exchanger stainless steel tube(mW)

H_g is the heat flowing through the helium exchange gas (mW)

P is the calculated pressure of the helium exchange gas (mW)

P_o is the observed pressure of the helium exchange gas (mW)

d is the mean free path in the gas (mm)

Table 5.3: **Adsorption of ^{4}He on Zeolite**

T_Z	P	V
4.2	3.8x10^{-4}	204
8.0	7.1x10^{-4}	41

where T_Z is the temperature of the zeolite

P is the pressure of the helium exchange gas (mbar)

V is the STP volume (cm^{-3}) of the gas pumped to P

ature of T_H=6.5K. On this basis we would expect the hot end low temperature limit to lie between 6.5K (Equation 5.1) and 8.5K (Equation 5.9), as indeed it does (see Section 5.4 and Figure 5.7). Similarly, when the inner vessel contains solid nitrogen (47K), the predicted on conductance is 70mW K^{-1}.

5.3.2 Zeolite Sorption Pump

For convenience the gas switch is filled to a pressure of one atmosphere at room temperature. The gas volume is 10 cm^3 at STP and from Table 5.1 we can see that for the switch to be turned off the pressure must be reduced to less than 2x10^{-5} mbar. Using adsorption isotherm data (Daunt and Rosen 1970, see Table 5.3), of ^{4}He on synthetic zeolite (Linde Molecular Sieve 13X) as a guide, the Z pump was filled with 1.5 g of zeolite pellets, dry density about 1.2 g cm^{-3}. Upon cooling to 5K the ^{4}He exchange gas was pumped to less than 10^{-5} mbar. A prototype VCGS demonstrated cold stage temperature control by warming the zeolite to static temperatures in the range 14–20K, resulting in exchange gas pressures of 10^{-4}–10^{-3} mbar.

The Z pump is thermally isolated from the heat exchanger and cold–face by mounting it on a thin–walled stainless steel tube which conveniently serves the dual role of exchanger pumping line. Z pump warming is by a wire wound heater and cooling through a thermal link to the cold end of the exchanger, nominally at the cold–face temperature. The conductance of the thermal link is a compromise between heat leak to the liquid helium and Z pump cycling time constant. The latter is proportional to Z

pump thermal mass which is consequently minimised. Torre and Chanin (1984) have noted that a good choice of material for the link is pure tin as its conductivity falls over the range 4–20K, whereas that of copper rises. Unfortunately our tin links failed to work as predicted, possibly because of purity problems. Copper wire was used instead.

In the first prototype VCGS the gas capillary tube connected directly to the Z pump. When operated in its variable conductance mode, it was found to suffer from spontaneous thermal oscillations. A simulated cold stage experienced 1– 2K undamped temperature oscillations with periods ranging from half a minute to several minutes depending upon the cold stage temperature. This phenomenon (see for example Rose-Innes 1973), is caused by sound waves in helium gas driven by large temperature gradients in tubes of certain geometries and is used to good effect in sensing the liquid helium surface in dewars ('dipping'). It was fully cured by connecting the capillary to the cold end of the heat exchanger (see Figure 5.3), effectively heat sinking the capillary to the liquid helium cold–face before entering the gas switch.

5.3.3 Electronic Temperature Controller

The purpose of the electronic temperature controller is to maintain the cold stage (and hence the detector) to a selectable temperature stability down to ± 0.1K over the range 10–60K and to automatically move between selected temperatures. It achieves this by providing power for the Z pump and cold stage heaters. The power (voltage) settings are selected by hard–wired logic depending upon the actual temperature and the target temperature set by the operator. The operator must also select the Z pump heater voltage (and therefore Z pump temperature) required for the chosen detector temperature, using a cold stage temperature versus Z pump voltage characteristic (see Figure 5.7). Examples of detector temperature control are illustrated in Figure 5.8.

5.3.4 VCGS Thermal Modelling

Thermal modelling of the VCGS is based on the thermal circuit drawn schematically in Figure 5.5. It was used to predict transient and steady state performance. Temperature nodes are linked by heat flow paths which are represented by thermal resistances (see Chapman 1984). Only the significant resistances are considered. The thermal resistance

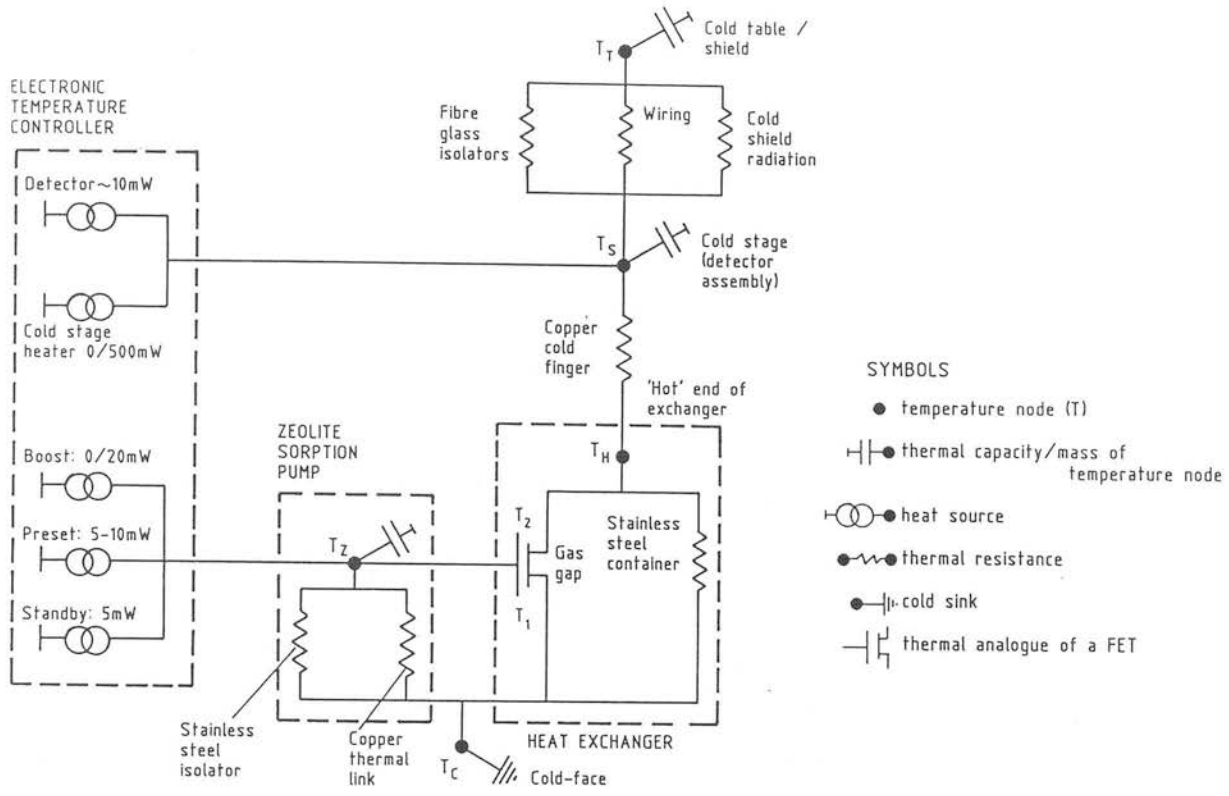


Figure 5.5 : Detector thermal model.

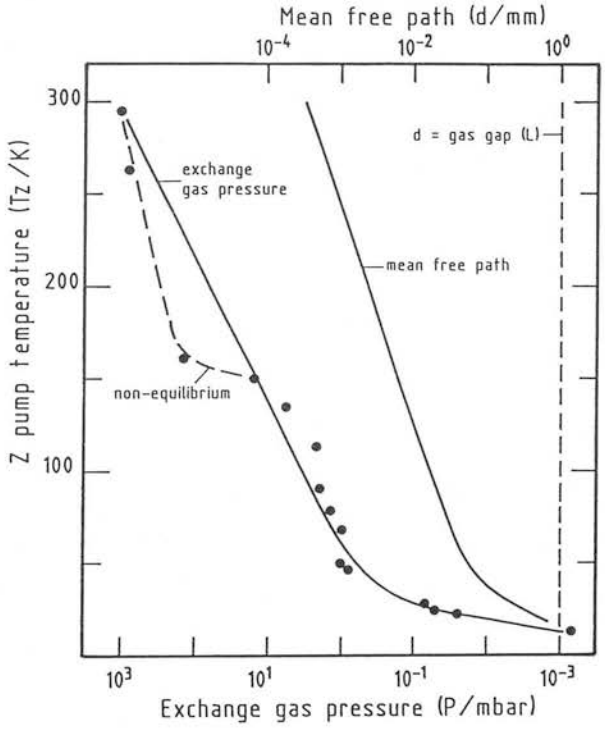


Figure 5.6 : Plot of Z–pump temperature against helium exchange gas pressure and mean free path. The gas switch is 'on' (ie. in its high thermal conductance state) when the mean free path $d \ll L$ (the distance separating the heat exchange cylinders). More recent measurements with a pressure gauge better calibrated at higher pressures, suggests that measured pressures $> 10^0$ mbar are greater than shown. However this indication only reinforces the requirement that $d \ll L$ in the 'on' state.

of each conduction path is proportional to L/A, where L is the path-length and A its cross-sectional area. In the case of the gas, L is the gas gap and A the common surface area of the exchange tubes. The thermal capacities of the temperature nodes and thermal conductances of the links are both dependent upon temperature. Initial modelling addressed the thermal resistance of the heat exchanger tubes and concluded that with copper tubes of 1.0 mm wall thickness, the thermal resistance of the exchange tubes was at most 20% of that across the gas gap. Radiative coupling of the hot and cold tubes is insignificant.

The VCGS thermal model can be improved by introducing more component nodes. However, in practice the model is limited by the accuracy to which the physical parameters are known. VCGS performance agrees with that predicted to within 20% in most cases. The exceptions to this are the on conductances of the heat exchanger when there is, respectively, solid nitrogen and liquid helium in the cryostat inner vessel. In these cases the measured conductances are 40% lower than predicted (see Section 5.4). However, this has not resulted in any significant loss of performance.

5.4 Implementation and Performance

Because of their thermal properties, electrolytic tough pitch (ETP) copper and stainless steel are used almost exclusively in construction of the gas switch. The exchange tubes are soldered into their respective hot and cold copper housings with high melting point solder. Both the heat exchanger and Z pump stainless steel tubes have had their wall thicknesses lathed down to 0.15mm in order to meet the thermal isolation requirements. They were assembled with low melting point solder. A modified KF10 sintered copper filter of the type found in vacuum gauges is used as a plug for the zeolite pellets. To allow the zeolite charge to be changed, a demountable indium wire seal interfaces the Z pump with its isolation post. Lake Shore Cryotronics DT500 silicon diode temperature sensors are mounted on both the Z pump and cold stage. The 500 Ω Z pump heater is wound with 6.3m of SWG 44 manganin wire. A $\frac{1}{16}$ inch bore stainless steel capillary connects the gas switch with a vacuum valve and Pirani pressure gauge mounted on the cryostat exterior. Swagelok ferrule and nut connectors vacuum seal the the capillary at both the hot and cold joints. To avoid undue strain on the vacuum-tight heat exchanger

the cold-finger is joined to it by flexible copper foils. The fully assembled VCGS, minus only its zeolite charge was chemically cleaned and then heated in a vacuum oven prior sealing.

The zeolite pellets were dried in a vacuum oven at 120°C for several days before being sealed into the Z pump. After leak–testing the gas switch was put through several helium purge and pump cycles before it was finally back–filled and valved–off.

A typical cooling curve for the cold stage is shown in Figure 4.12 and 4.13. Operating temperature is reached within 18 hours. Figure 5.6 shows the relationship between Z pump temperature and helium exchange gas pressure as measured at the Pirani gauge (nitrogen calibrated and corrected for helium). Although the effective gas temperature is not necessarily equal to the Z pump temperature, it is not significantly different above about 30K. On this basis the mfp in the heat exchanger is also plotted. Above 30K the mfp is orders of magnitude less than the gas gap, keeping the gas switch on, as required during cooling. In fact the data points at 50K and 46K are with solid nitrogen in the inner vessel (effective gas temperature and Z pump temperature the same). With a cold–face temperature of 46.1K, the cold stage reached 48.1K, while the hot end of the exchanger and Z pump stabilised at 46.8K and 46.1K respectively, resulting in a heat exchanger on conductance of about 45mW K^{-1}.

The cold stage temperature (T_S) versus Z pump temperature (T_Z) characteristic is plotted in Figure 5.7. The variable conductance regime is the linear region from T_S=10K (T_Z=17K) to T_S=60K (T_Z=13K). At T_S=10K the characteristic 'turns over', as expected, since the helium exchange gas pressure increases to the point where the mfp becomes less than the gas gap. Increasing T_Z above 20K does not significantly reduce T_S, which reaches a limit of about 9K, since conductance has become independent of gas pressure. Due to the 1.5K temperature gradient in the cold–finger, the hot end of the heat exchanger reaches the lower temperature of 7.5K, and with the liquid helium cold–face at 5.0K, the heat exchanger on conductance is 15mW K^{-1}. This compares with an off conductance (T_S=60K) of less than 0.5mW K^{-1}. Allowing for thermal transpiration corrections, the observed helium exchange gas pressures compare favourably with those predicted (see Table 5.2), except at the lower pressures where the pressure gauge 'bottomed out'.

From the variable conductance region of the T_S versus T_Z (and Z pump voltage V_Z)

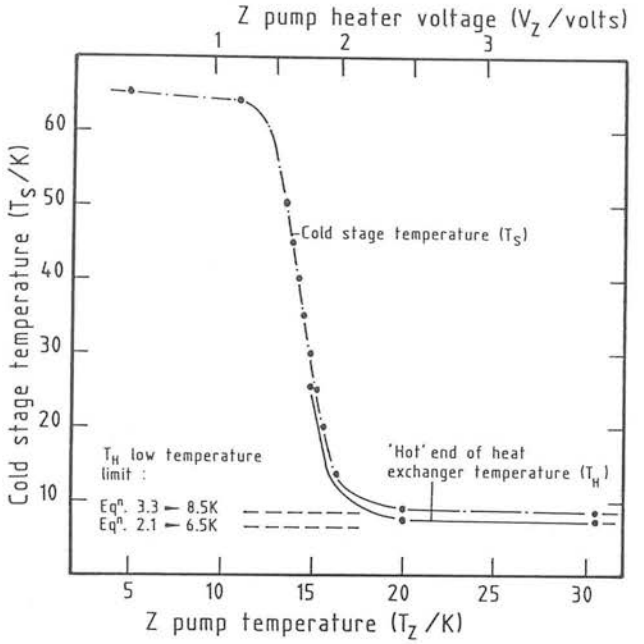


Figure 5.7 : Plot of the cold stage (detector assembly) temperature against Z–pump temperature and Z–pump heater voltage.

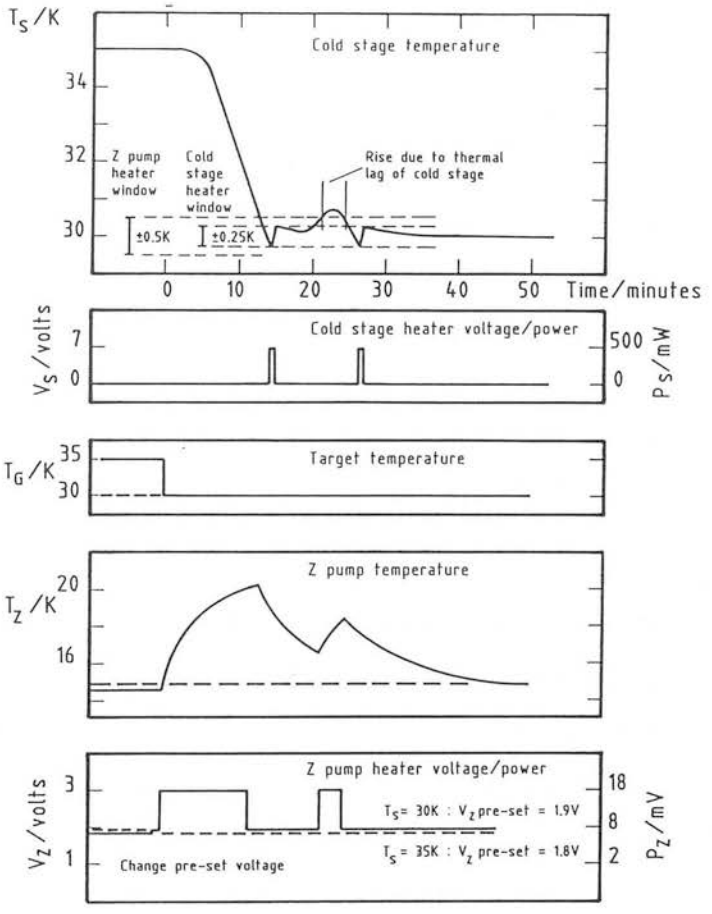


Figure 5.8 : An example of cooling the cold stage from 35K to 30K.

characteristic (Figure 5.7), it can be seen that to maintain T_S to $\pm$0.1K, T_Z must be stable to $\pm$0.01K and V_Z to $\pm$1mV. Such a level of control requires feed-back control of the Z pump. By setting both the Z pump and cold stage heater control windows to $\pm$0.1K (see Figure 5.8), the temperature controller can maintain this stability without controlling T_Z to $\pm$0.01K, but because of the long Z pump time constant excessive liquid helium usage may result. The best method is to have the Z pump and cold stage heater windows set at $\pm$0.5K and $\pm$0.1K respectively, while keeping the Z pump slightly warmer than optimum (about 0.2K). Since the exchanger conductance is now very slightly higher than required, the controller will maintain the cold stage temperature by slightly more frequent cycling of the heater. This method of control increases liquid helium consumption by only 10% above optimum. Naturally, things are easier if the stability requirements are relaxed.

When the Z pump is allowed to cool below about 10K it takes about 20 minutes of heating at 20mW before gas is desorbed by the zeolite. This is due to poor thermal coupling of the copper cell to the zeolite pellets when the helium exchange gas is completely adsorbed. Since effectively all the helium exchange gas is adsorbed at about 12K, the temperature controller prevents the Z pump from cooling below 12K by providing a "standby" heat input of 5mW (V_Z=1.5 volts). As explained above, the heat input to the Z pump required for variable temperature control is a compromise between heat leak to the liquid helium and Z pump cycling time. Using an ETP copper (AWG 30) wire thermal link of length 60mm and diameter 0.25mm (A/L=8x10^{-4}mm), connected between the Z pump and cold end of the heat exchanger (nominally the cold-face), a heat leak to the liquid helium of only 10mW is needed to achieve a detector temperature range of 10–60K (see Figures 5.7 and 5.8). This is an order of magnitude less than the heat leak required in the fixed link approach.

The reason for the sensitivity of T_S to changes in T_Z is probably due to the pressures required to operate in the variable conductance regime. To make T_S less sensitive we need to work at higher pressures, since then relatively more gas must be adsorbed or desorbed to change T_S, which requires a correspondingly greater change in T_Z. (Since the helium exchange gas is saturated, increasing the volume available to the gas will not help). Therefore, since the exchange gas pressure is inversely proportional to the surface area of the exchange tubes, this area must be decreased. But to maintain the low temperature limit and 'on' conductance, the gas gap must be reduced, presenting

potential construction problems.

After a temperature cycle of the IRCAM cryostat, the T_S versus T_Z characteristic is sometimes found to shift in T_Z by up to ± 0.3K. Possible causes for this effect include some form of zeolite hysteresis, different heat loads from the detector thermal support structure, and gas redistribution between the switch and Pirani gauge. The slope does not change and so in such cases the temperature controller is just 'retuned'.

The upper temperature limit of the VCGS may be extended by using the cold stage heater. To reach 80K a 10mW heat input is required. If a permanent operating temperature is decided upon, an optimised thermal link may be placed in parallel with the heat exchanger. This avoids using the Z pump and its slight heat load penalty. A thermal link in parallel with the heat exchanger, lowers the low temperature limit, however, for a given cold stage heat input, it also lowers the upper temperature limit. Performance gains in the VCGS design may still be made by reducing the gas gap, while its construction can be improved by replacing the capillary and indium seal with cold welds, making the VCGS completely self contained. In the current VCGS the on conductance is limited largely by the width of the gas gap. When care is taken to minimise the thermal resistance of the heat exchanger tubes, the thermal conductance of the heat exchanger is given by Equation 5.2, and is consequently inversely proportional to the gas gap.

Chapter 6

Near–IR Imaging and Imaging Polarimetry of OMC 2

6.1 Introduction

Some $12\,\hat{'}$ north of the Trapezium OB star cluster and the M42 HII region in Orion, lies a cluster of 5 or so compact near–IR sources with extended emission, about $2\,\hat{'}$ in diameter, known as Orion Molecular Cloud 2 (OMC 2). The cluster is associated with a peak in molecular gas and dust emission which forms the central ($10\,\hat{'}$x $40\,\hat{'}$) north–south ridge of the huge dark cloud L1641 (Goudis 1982). OMC 1 is another molecular cloud associated with the more luminous $0.5\,\hat{'}$ diameter Kleinmann–Low compact IR cluster situated $1\,\hat{'}$ north–west of the Trapezium. Together with a third cloud $16\,\hat{'}$ north of OMC 2 (OMC 3), these regions maybe the result of a fragmentation process which occurred in L1641. Following discovery of OMC 2 by virtue of its IR and strong CO emission by Gatley *et al.* (1974), the region has been extensively observed in both the continuum and in molecular emission lines. Continuum observations include 1 μm photography (Cohen and Frogel 1977), 40–400 μm far–IR photometry (Thronson *et al.* 1978), and submillimetre maps at 400 μm (Smith *et al.* 1979) and 1000 μm (Westbrook *et al.* 1976). The molecular emission line observations include the vibrationally excited H_2 lines at 2 μm (Fischer *et al.* 1980, Thronson and Thompson 1982), HCN (Morris *et al.* 1974),NH_3 (Batrla *et al.* 1983),H_2CO and CO (Kutner *et al.* 1976), and velocity resolved CO (Fischer *et al.* 1985). Most recently, OMC 2 has been the subject

of an extensive programme of near–IR (1–5 μm) photometry and polarimetry, together with mid– to far–IR (5–100 μm) photometry by Pendleton *et al.* (1986).

The absence of optical counterparts, the presence of an infrared excess, and the close relationship to molecular and dust emission peaks, strongly implies that star formation is occurring in OMC 2. The far–IR observations of Thronson *et al.* (1978) show OMC 2 to be similar in certain respects to OMC 1. OMC 2 is about 100 times less luminous, 4 times less massive and has average densities about 4 times smaller; consistent with a young cluster of B or A stars. The cluster has a far–IR luminosity of about 2 x 10^3 $L_\odot$ into a 3.5$\hat{}$ beam, most of which is attributed to emission from cool dust (T_d=30K) within about 0.1 pc. (50$\hat{''}$ beam diameter at 500 pc.) of the near–IR source OMC 2 Irs4. Using 4$\hat{''}$ beams Pendleton *et al.* were able to resolve Irs4 into two components, Irs4 South and Irs4 North, with a separation of 4$\hat{''}$. From 1–30 μm photometry, Pendleton *et al.* reason that most of the longer wavelength flux originates from a point centred on Irs4 N. By associating the 50 μm and 100 μm flux with Irs4 N, Pendleton *et al.* conclude that Irs4 N is the dominant luminosity source in OMC 2, being 3 times brighter than the other significant luminosity source OMC 2 Irs1. Assuming that Irs1 and Irs4 make up the total luminosity of OMC 2, then $L_{Irs1} \leq 500\,L_\odot$ and $L_{Irs4} \leq 1500\,L_\odot$. From 3.8 μm polarimetry Pendleton *et al.* find that the near–IR flux from Irs4 N is best explained by scattering from Irs4 S and is not intrinsic. Given that Irs4 N is intrinsically bright at longer wavelengths, they tentatively suggest a model in which the 3.8 μm flux is due to scattering from an extended envelope or dust shell surrounding Irs4 N and that Irs4 S is the illuminating source. Irs 4 is also closely associated with 2 μm H_2 line emission (Fischer *et al.* 1980, Thronson and Thompson 1982), and the overlapping blue– and redshifted lobes of a high velocity CO flow lie only 15$\hat{''}$ to the north. Further evidence of mass outflow from a young stellar object (YSO) comes from the discovery of an H_2O maser at the position of Irs4 (Genzel and Downes 1979). The extended 2.2 μm and 3.8 μm emission to the north–east of Irs1 is shown by Pendleton *et al.* to be a good example of a near–IR reflection nebula illuminated by Irs1, while the flux from Irs1 itself is perhaps best characterised as reddened emission from a 1000K dust envelope surrounding a protostar. Because of the compact nature of OMC 2, higher spatial resolution would be beneficial in further characterising the evolutionary status of the region.

Using a 1–5 μm infrared camera we have been able to obtain seeing–limited multi–

colour images as well as 2.2 μm imaging polarimetry of all the compact near–IR sources in OMC 2. Our results imply that OMC 2 Irs1 and OMC 2 Irs4 N each illuminate bipolar reflection nebulae and OMC 2 Irs2 is found to illuminate a monopolar reflection nebula. Multi–colour observations of the nebulae have enabled us to deredden the near–IR flux from these objects, from which we conclude they are low–to–intermediate mass pre–main sequence stars embedded in circumstellar disks. The orientation of the disks does not appear to be related to the general magnetic field in the region.

6.2 Observations

6.2.1 Instrumentation and Procedure

The observations presented here were made at the 3.8m UK Infrared Telescope on Mauna Kea, Hawaii, during January and February 1987, while commissioning the 1–5 μm infrared camera, IRCAM (McLean *et al.* 1986, McLean 1987). Follow–up observations were made in December 1987. Images taken during January and February 1987 were made with an engineering–grade array (FPA#004) having a high quantum efficiency (>60%, see Chapter 2.3), low read–out noise (500 electrons RMS), and a well–depth of 10^6 photo–electrons. A photo–emission effect caused by faulty MOSFETS on the silicon multiplexer , resulted in a high (non–thermal) dark signal of about 2 x 10^4 electrons per second. This array also had about 200 bad pixels (5% of total), the majority of which were located in a diagonal crack–like feature in the InSb slab. The December 1987 images were made with a science–grade array (FPA#061) having a dark current of only 200 es^{-1} and about 100 bad pixels. The images discussed here were obtained in STARE mode. This procedure involves taking three separate images of equal integration time (on–chip integration time multiplied by the number of co–adds): the object image (O), an image of a nearby area of blank sky (S), and an image of a cold dark slide or "blank" in the filter wheel (the dark current image, D). Following subtraction of the dark signal from each image, the object image is then flat–fielded by dividing by the corrected sky image. This image is then re–scaled to the mean of a 'clean' area in the sky image:

$$\frac{(O-D)}{(S-D)}\overline{(S-D)}_{clean} \tag{6.1}$$

Data lost due to bad pixels (including cosmic ray hits) is replaced by data from additional images offset in RA and Dec from the original set. Any remaining bad pixels are replaced by interpolating data from surrounding pixels. Offset images also enable the identification of image ghosts. By using a combination of flux standards and previous observations of OMC 2, the absolute photometric accuracy in these images is found to be better than 20%. Relative (on–chip) photometry is better than 5%. The data have not been corrected for air mass since only small changes (5–10%) in the detected flux are seen between 1 and 2 airmasses (McCaughrean 1988).

Polarimetry is accomplished by imaging through a warm achromatic wave–plate external to IRCAM in conjunction with a cold filter and cold wire–grid polariser inside IRCAM. The polarisation module (IRPOL) is placed in the f/36 converging beam prior to all oblique reflections. This arrangement minimises the instrumental polarisation. The wave–plate used here had a half–wave retardance between 1 μm and 3 μm. Images ($I(\theta)$) of the object, and an associated sky flat,with the wave–plate positioned at 0°, 45°, 22.5° and then 67.5° from datum, are taken consecutively. Any offset of the datum from the equatorial system is derived from observations of objects with known polarisation position angles. The images taken through each wave–plate position are reduced in the same manner as unpolarised images. The unnormalised Stokes parameters (I, Q, U) for each pixel (x,y) are derived from :

$$I(x,y) = I_0(x,y) + I_{45}(x,y) \text{ and } I(x,y) = I_{22.5}(x,y) + I_{67.5}(x,y) \tag{6.2}$$

$$\frac{Q(x,y)}{I(x,y)} = \frac{I_0(x,y) - I_{45}(x,y)}{I_0(x,y) + I_{45}(x,y)} \tag{6.3}$$

$$\frac{U(x,y)}{I(x,y)} = \frac{I_{22.5}(x,y) - I_{67.5}(x,y)}{I_{22.5}(x,y) + I_{67.5}(x,y)} \tag{6.4}$$

where I_0, $I_{22.5}$, I_{45} and $I_{67.5}$ are the reduced images at the respective wave–plate positions. The (fractional) linear polarisation image is given by:

$$P(x,y) = \frac{[U^2(x,y) + Q^2(x,y)]^{\frac{1}{2}}}{I(x,y)} \tag{6.5}$$

and the position angle of the electric vector is given by :

$$\phi(x,y) = \frac{1}{2}\tan^{-1}\frac{U(x,y)}{Q(x,y)} \tag{6.6}$$

Either the P and ϕ images, or the Q and U images, may be combined to produce a polarisation vector map which can be overlaid on the total intensity image I. To enable

comparison with polarimetry gathered with aperture photometers, it is a simple matter to place software apertures of the correct size at the required locations in each of the four wave–plate position images (I_θ), and then re–calculate P and ϕ. In order to avoid spurious polarisation through mis–matching, it is important that the four wave–plate images be carefully registered, ideally on an unpolarised point source common to each image. This is more of a problem at the 2.5″/pixel scale since point sources are undersampled and precise registration is therefore not possible, resulting in spurious polarisations on the point sources. Similarly, as the images are not acquired simultaneously, good (steady) seeing and constant atmospheric transmission are required for accurate polarimetry. Due to the size of the wave–plate, all but the central arc–minute of the 2.5″/pixel scale Stokes images are occulted. Observing conditions were excellent. The Stokes images have been calibrated by observing the Becklin–Neugebauer (BN) object and by comparison with previous polarimetry, such as OMC 2 Irs3 (Dyck and Londsdale 1979) and the OMC 2 reflection nebulae observed by Pendleton *et al.* We have also obtained an undersampled 20 point 2.2 μm polarisation map of the OMC 2 Irs1 region using the UKT 6 single–channel InSb system together with the Kyoto polarimeter, employing a 12.6″ aperture. This map provides us with an independent calibration source. Instrumental polarisation is found to be less than about 0.2%.

6.2.2 Results

The images acquired are listed in Table 6.1. The limiting surface brightness has been assessed by examining the base level noise and comparing this with aperture photometry on the faintest detectable extended emission. Both methods give the same 3σ limits. Significant noise components include those due to dark current subtraction, flat–field division and the read–noise. However, an additional noise component due to an intermittent problem with the engineering array, means that it is not possible to properly extrapolate the surface brightness limits between some of the images. Four filters have been used : broad band J (1.25 μm,$\Delta\lambda$ = 0.3 μm), H (1.65 μm,$\Delta\lambda$ = 0.35 μm) and K (2.2 μm,$\Delta\lambda$ = 0.42 μm), and narrow band L (nbL, 3.6 μm,$\Delta\lambda$ = 0.07 μm). Both the J and H filters used at the time were found to suffer from long wavelength leaks, but we believe this has not significantly affected the data since, with the exception of Irs4 N, the faint J objects are not particularly red. Due to the strong thermal backgrounds

longward of 3 μm, the large photon–gathering capability of the 3.8m UKIRT, and the limited charge–storage capacity of the detector, it is not possible to integrate in the broad band L′ filter (3.8 μm,$\Delta\lambda$ = 0.67 μm) without saturating , and so the complementary narrow band filter was used. During the high–resolution imaging sessions the seeing conditions resulted in point–source FWHM of $\leq 2''$. Due to an RA oscillation effect in the telescope, the point–sources have a slightly elliptical appearance.

The dynamic range (maximum signal divided by base noise level) of these images is about 1000, while that of the video display unit used is 256. Consequently it is necessary to display an image at more than one contrast level to see all the possible structure. Figures 6.1 and 6.2 show the low–resolution (2.5 $''$/pixel) J and K intensity images centred on the near–IR cluster in OMC 2, together with their contrast stretched versions. At this image scale parts of the collimator cause vignetting in the corners of the array. Data in these areas has reduced S/N but is still useable. By centroiding on the positions (corrected for proper motion) of the Parenago stars (Parenago 1954) identified in Figure 6.1(a), we determine an image scale of $2.48 \pm 0.05''$. The error on this scale results in a $\pm 3''$ relative position error across the images. However, accuracies of $\pm 1''$ can be obtained by offsetting from Parenago stars within the images. The RA and declination scale of Figures 6.1(a) and 6.2(a) are offset from the point–source OMC 2 Irs3 which we find to have coordinates RA (1950) $5^h\ 32^m\ 59.1 \pm 0.1^{sec}$, Dec $-5°\ 12'\ 10 \pm 1''$, in agreement with the determination by Gatley *et al.* (1974). From a 126 image mosaic of the near–IR cluster in OMC 1, the high–resolution image scale is found to be $0.625 \pm 0.001''$/pixel. The corresponding fields of view are $154''$ x $144''$ (E–W x N–S) and $39''$ x $36''$ at low and high–resolution respectively. The low–resolution scale has an unvignetted circular FOV, $135''$ in diameter (see Chapter 3.2.1).

Since faint point–sources subtend one pixel in the low–resolution images, it is necessary to take one or more offset images in order to eliminate spurious or noisy pixels. Even then it is not simple to avoid mis–identifications at the fainter levels. For this reason we do not identify single–pixel events unless they appear in more than one colour. Software aperture photometry of the point–sources and extended emission is given in Tables 6.2 and 6.3. Magnitudes are found to be aperture dependent, but due to the crowded nature of the field and ubiquity of extended emission, it is not possible to use apertures significantly larger than $12''$ in diameter on the low–resolution images. Similarly, $4''$ and $3''$ diameter apertures are used on the high–resolution images. Wherever

Table 6.1: **OMC 2 Data Log**

Object	*Filter*	*Scale*[a]	*Exp.*[b]	*3σ limit*[c]	*UT Date*	*Air mass*
OMC 2	J	2.5	200	18	1 Jan 87	1.15
OMC 2	H	2.5	400	17	9 Feb 87	1.90
OMC 2	K	2.5	500	17	1 Jan 87	1.20
Irs2/4	J	0.63	75	18	25 Jan 87	1.9
Irs2/4	H	0.63	150	16	25 Jan 87	1.8
Irs2/4	K	0.63	150	16	25 Jan 87	1.6
Irs2/4	nbL	0.63	90	11[d]	25 Jan 87	1.7
Irs2/4	H	0.63	300	19	31 Dec 87	1.14
Irs2/4	K	0.63	300	17.5	31 Dec 87	1.25
Irs2/4	nbL	0.63	300	11.5	31 Dec 87	1.16
Irs1	H	0.63	45	16.5	25 Jan 87	1.45
Irs1	K	0.63	45	16.5	25 Jan 87	1.33
Irs1	nbL	0.63	45	11	25 Jan 87	1.50
OMC 2	K+pol	2.5	300[e]		7 Feb 87	1.5
Irs2/4	K+pol	0.63	150[e]		26 Jan 87	1.8
Irs1 nebula	K+pol	0.63	300[e]		26 Jan 87	1.6

[a] Image scale in arcseconds per pixel.

[b] On source integration time in seconds.

[c] 3σ surface brightness limit in magnitudes per sq. arcsecond.

[d] Transformed into L of the Johnson system.

[e] Per waveplate position.

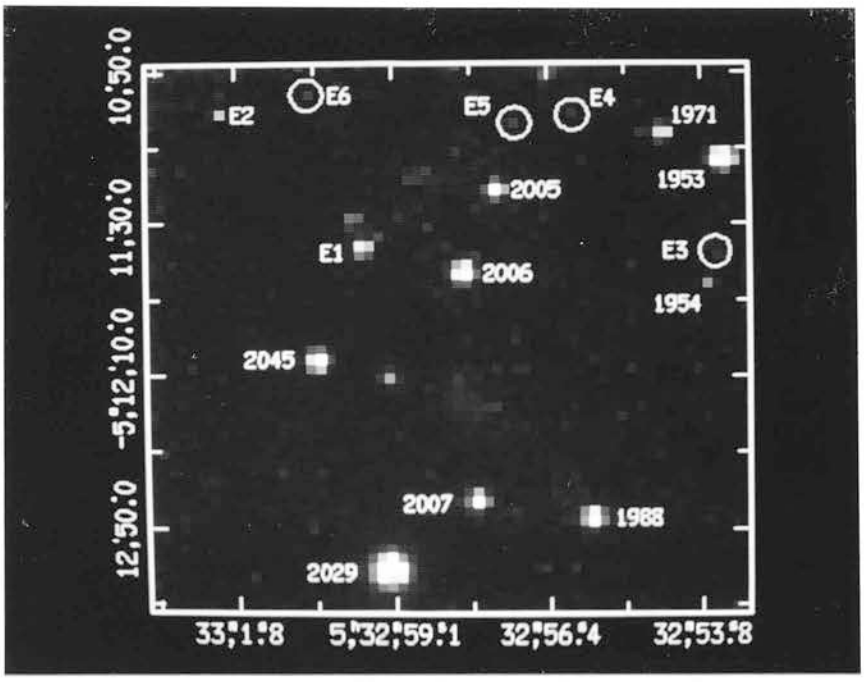


Figure 6.1(a) : Low–resolution ($2.5\,\hat{''}$ pixel^{-1}) J ($1.25\,\mu$m) image of OMC 2. The non–IR sources are identified.

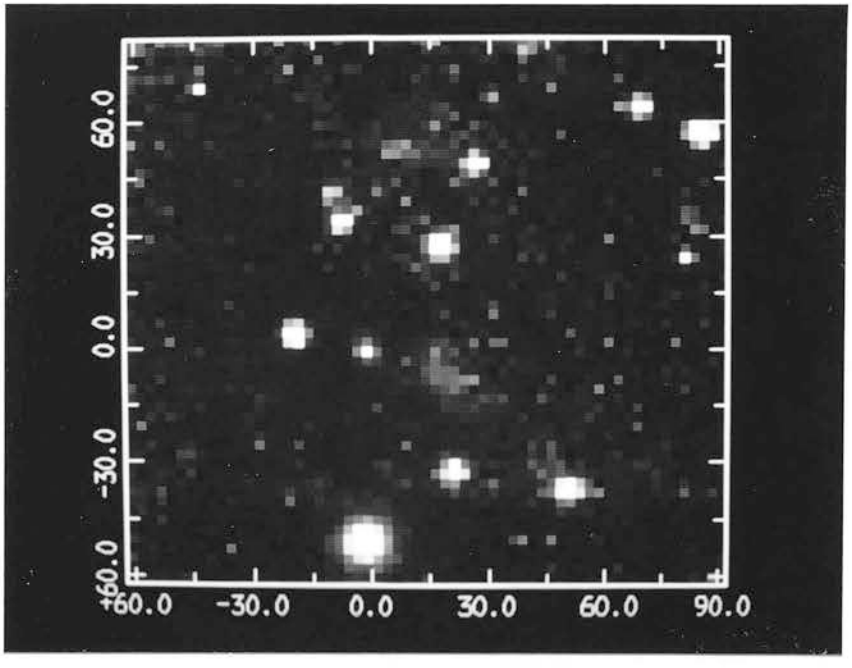


Figure 6.1(b) : Low–resolution ($2.5\,\hat{''}$ pixel^{-1}) J ($1.25\,\mu$m) image of OMC 2. The contrast enhanced version of Figure 6.1(a), showing the fainter nebulosity (3σ limiting surface brightness of 18 magnitudes per sq. arcsecond). The offsets are in arcseconds from Irs3 (see Figure 6.2(a)).

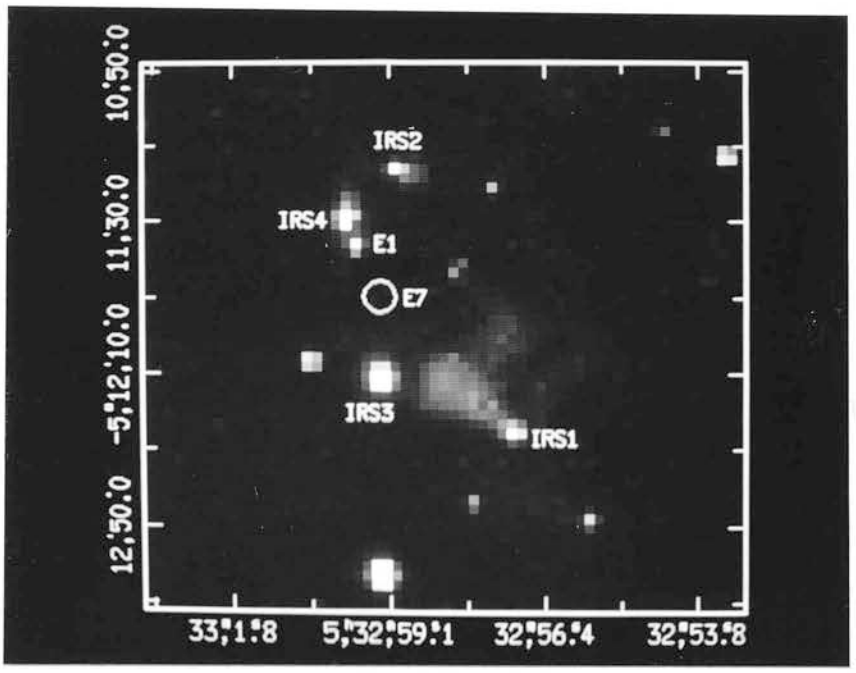


Figure 6.2(a) : Low–resolution ($2.5\,\widehat{''}$ pixel^{-1}) K ($2.2\,\mu$m) image of OMC 2. The near–IR sources are identified.

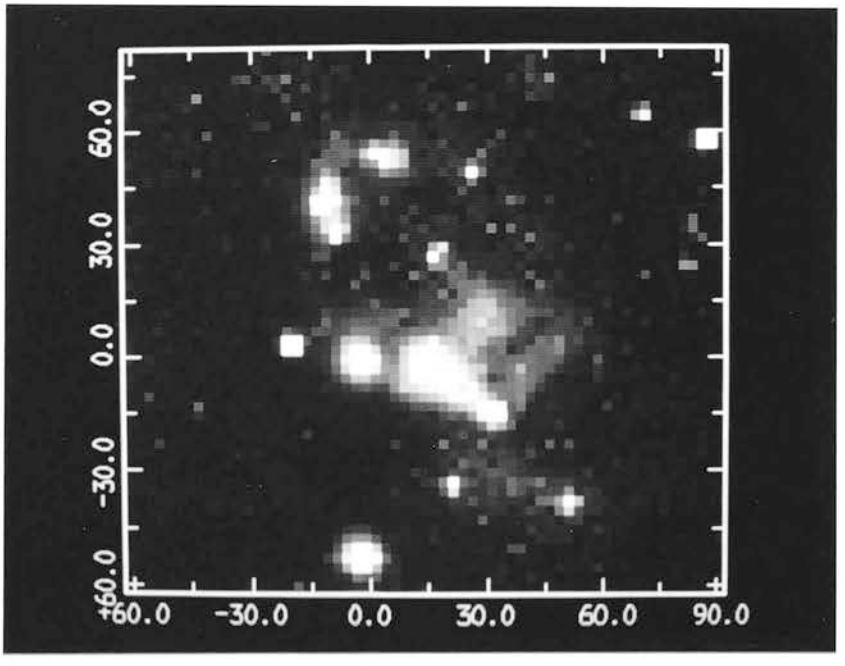


Figure 6.2(b) : Low–resolution ($2.5\,\widehat{''}$ pixel^{-1}) K ($2.2\,\mu$m) image of OMC 2. The contrast enhanced version of Figure 6.2(a), showing the fainter nebulosity (3σ limiting surface brightness of 17 magnitudes per sq. arcsecond). The offsets are in arcseconds from Irs3.

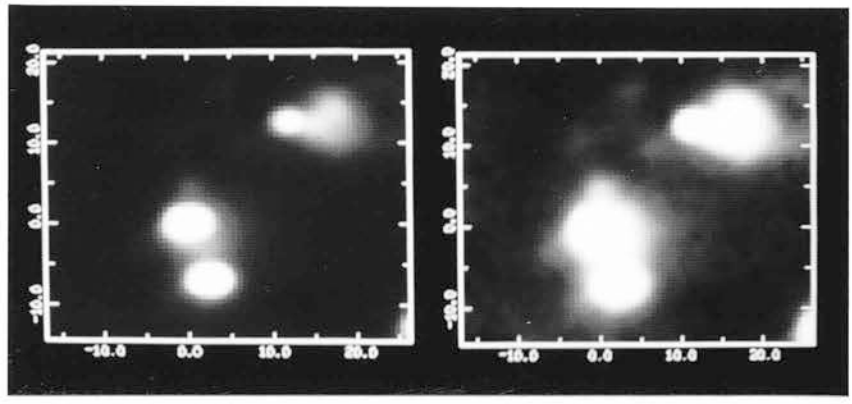


Figure 6.3(a) : High–resolution ($0.6\,\widehat{''}$ pixel^{-1}) H ($1.65\,\mu$m) image of OMC 2 Irs2/4. Low and high contrast versions.

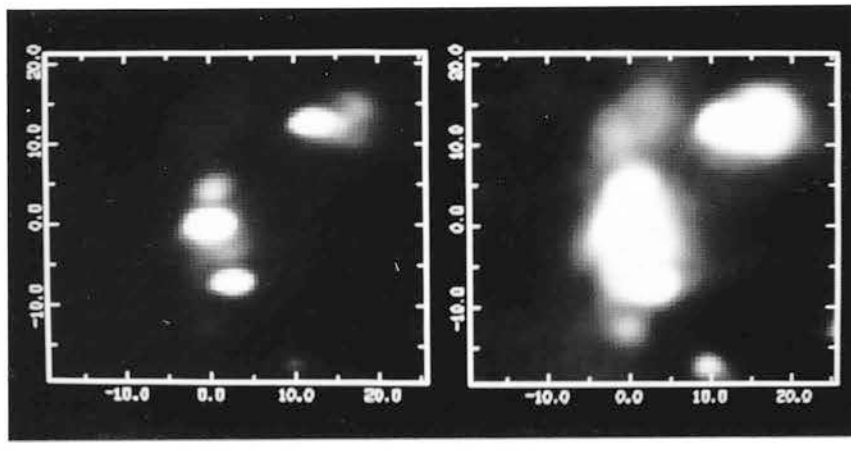


Figure 6.3(b) : High–resolution ($0.6\,\widehat{''}$ pixel^{-1}) K ($2.2\,\mu$m) image of OMC 2 Irs2/4. Low and high contrast versions.

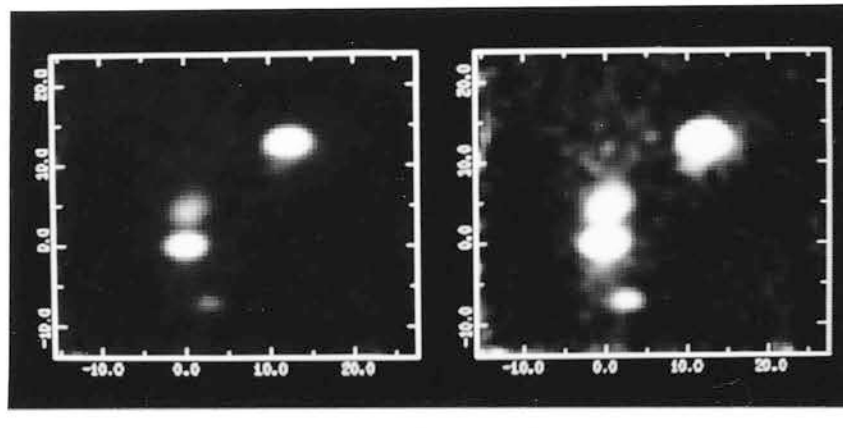


Figure 6.3(c) : High–resolution ($0.6\,\widehat{''}$ pixel^{-1}) narrow band L ($3.6\,\mu$m) image of OMC 2 Irs2/4. Low and high contrast versions.

Irs2 is the source towards the top–right (north–west) of each image. The double source Irs4 (North and South) lies about $20\,\widehat{''}$ south–east of Irs2. E1 is the relatively blue object lying about $8\,\widehat{''}$ south of Irs4 S. The offsets are in arcseconds from Irs4 S.

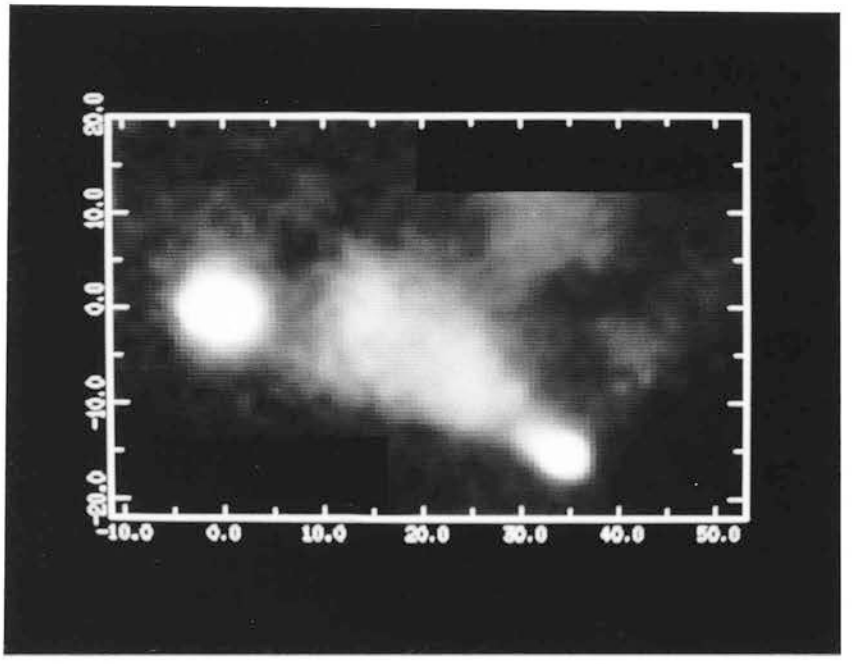


Figure 6.4 : High–resolution ($0.6\,\widehat{''}$ pixel^{-1}) two image K ($2.2\,\mu$m) mosaic of the OMC 2 Irs1/3 region. The bright extended emission is nebulosity illuminated by Irs1 which is located in the bottom–right (south–west) of the image. The offsets are in arcseconds from Irs3, the bright source in the east of the image.

it has been possible to vary the aperture size, the affects on the measured colours have been minor. The low–resolution J and K images were used to select regions for multi–colour imaging at higher resolution and for imaging polarimetry. The H,K and nbL images centred on the OMC 2 Irs2/4 region are shown in Figure 6.3. Figure 6.4 is a mosaic (of two frames) at K of Irs1 and the reflection nebula to the NE. Table 6.4 lists the results of software aperture polarimetry obtained from the Stokes images centred on Irs2/4 and the Irs1 nebula.

6.3 Discussion

The position of all the point–sources are identified in Figures 6.1(a) and 6.2(a) and the photometry is summarised in Table 6.2. Nine of the sources are field stars previously catalogued by Parenago (1954) and are given the prefix π. Gatley *et al.* (1974) first identified the compact near–IR sources OMC 2 Irs1–5, however Irs5 is also π2045. On the basis of its J–K and H–K colours, Irs5/π2045 is not a near–IR source. We have identified 7 previously uncatalogued sources, which are given the prefix E (étoile), and which are listed in order of increasing J magnitude. With the exception of E1 and E7 all other E sources are only imaged at low–resolution (2.5″/pixel) and so cannot strictly be classed as point sources. E1, E2, E4 and E5 can be seen in the 1 μm image of Cohen and Frogel (1977); E1 is discussed further in the section on OMC 2 Irs4. E2, E3 and E5 appear to be faint field stars, E7 is only seen in the low– and high–resolution K images. E4 and E6 are both associated with 2.2 μm nebulosity, which possibly contaminates their J–K colours. E6 lies about 10″ east and 40″ north of OMC 2 Irs4 in a region of molecular H_2 S(1) v=1–0 line emission thought to be fluorescence (Burton priv. comm.), and its red colour of J–K=1.7±0.4 makes it a potential excitation source for the emission. The field stars are clearly separated from the near–IR sources on the basis of their bluer colours (J–K, H–K, see Table 6.2).

6.3.1 OMC 2 Irs2 Region

The compact source Irs2 is not visible at J, however the extended emission associated with it is clearly seen in both the high– and low–resolution images. At H, K and nbL, Irs2 is revealed as a point source with magnitude decreasing with wavelength (see

Table 6.2: **OMC 2 Point Source Photometry**

Object	*Beam*[a]	*Magnitude*[b]				*Colour*[c]		
		J	H	K	L	J-K	H-K	K-L
π1953	12	10.2		9.4		0.7		
π1954	12	12.8		11.7		1.1		
π1971	12	11.6		11.1		0.5		
π1988	12	10.8	9.9	10.0		0.9	-0.1	
π2005	12	11.4	10.5	10.8		0.6	-0.3	
π2006	12	10.5	10.5	10.5		0.0	0.0	
π2007	12	11.2	10.5	10.4		0.7	0.1	
π2029	12	8.9	7.9	7.7		1.3	0.2	
π2045[d]	12	10.9	9.7	9.7		1.2	0.0	
E1	12	11.6	9.9	9.4		2.2	0.5	
E1	4	12.0	10.7	10.2	9.5	1.8	0.5	0.7
E2	12	12.8		12.0		0.8		
E3	12	12.9		11.6		1.3		
E4	12	13.3±0.2		11.5		1.8		
E5	12	13.4±0.2		12.3±0.3		1.1±0.4		
E6	12	13.5±0.3		11.8±0.2		1.7±0.4		
E7	12			13.7±0.3				
E7	3			13.4±0.2				
Irs1	4	>14.8	12.1	9.5	6.7	>5.3	2.6	2.8
Irs2	4	>14.1	12.2	9.9	6.9	>4.1	2.3	3.0
Irs3	12	12.6	9.4	7.7		4.9	1.7	
Irs4 S	4	13.2	10.5	8.9	7.5	4.3	1.6	1.4
Irs4 N[e]	4	> 15.5	14.0±0.2	11.3	8.5	> 4.6	2.7	2.8
Irs4 N	3	> 16.5	14.7±0.3	11.7	8.9	> 5.2	2.5±0.3	2.9
Irs4 N[f]	3	> 17.5	15.0±0.4	11.9	9.0	> 5.7	3.1±0.6	2.9

[a] Software aperture in arcseconds. The 12$\widehat{''}$ apertures are from the low–resolution images, while the 4$\widehat{''}$ and 3$\widehat{''}$ apertures are from the high–resolution images. [b] ±0.1 unless stated. No entry means no coverage or no identification. [c] ±0.2 unless stated. [d] OMC 2 Irs5. [e] Contaminated by Irs4 S. [f] Using sky near Irs4 S for subtraction, to minimise contamination.

Table 6.3: **OMC 2 Extended Emission Photometry**

Object	*Magnitude*[a]*/sq. arcsec*			*Colour*[b]	
	J	H	K	J-K	H-K
Irs1 (15,13)[c]	17.7	15.0	14.0	2.7	1.0
Irs1 (3,25)	18.7[d]	16.4	14.8	3.9	1.6
Irs1 (-14,17)	> 18.7[e]	17.2	15.6	>3.1	1.6
Irs1 (-13,-15)	18.1	17.0	15.8	2.3	1.2
Irs1 (-4, -18)	18.2	17.1	15.8	2.4	1.3
Irs2 (-6,2)	16.9	15.4	14.4	2.5	1.0
Irs4S (0,8)	> 18.7	18.3±0.2	15.8±0.2	>1.8	2.5±0.3

[a] ±0.1 unless stated.

[b] ±0.2 unless stated.

[c] Offset position : (15″E(+),13″N(+)).

[d] 2σ detection only.

[e] Not detected to the 2σ level.

Table 6.2). Pendleton *et al.* find the magnitudes of Irs2 to be H=13.0 and K=10.6 into a 4″ beam, from observations made in 1983. Our January 1987 results show Irs2 to be significantly brighter at H=12.2 and K=10.0. Follow–up images made in December 1987 are not significantly different to those taken in the previous January and so any possible variability is not confirmed. It is interesting to note that the apparent change in colour over the 4 years is parallel to both the reddening vector and locus of blackbody temperature, as would be expected if the changes were real and due either to material being removed from the line of sight, or to an increase in photospheric temperature . Morphologically, the extended emission is similar at J, H and K, and extends to about 10″ west of Irs2, with a maximum north–south extent of about 5″. Distinct structure exists within the nebula. A loop of brighter emission is seen in both the H and K images stretching from Irs2 to a knot 6″W and 2″N. In the contrast–stretched versions of the H and K images, two arc–like extensions to the nebula are seen at 12″W and 2″ above and below the east–west line of the nebula. These arcs are similar to features seen in GL2591 and which have been interpreted as limb–brightened edges of stellar cavities

Table 6.4: **OMC 2 Aperture Polarimetry at 2.2 μm**

Object	*Beam*[a]	$P/\%$	ϕ^b
Irs1 (8,6)[c]	4	22.5±0.9	144±2
Irs1 (14,9)	4	33.3±0.6	143±2
Irs1 (21,11)	4	27.0±1.1	145±2
Irs1 (3,25)	4	33.7±1.5	110±3
Irs2	4	9.8±1.2	3±5
Irs2	3	12.7±1.1	3±2
Irs2 (-6,0)	4	14.2±1.6	-2±4
Irs2 (-6,0)	3	15.2±1.3	2±3
Irs3	4	3.59±0.13	108±3
Irs3	3	3.5±0.3	100±5
Irs4 S	4	2.8±0.5	176±6
Irs4 S	3	2.8±1.0	5±13
Irs4 S (1,8)	3	30±3	102±5
Irs4 N	4	20.0±1.4	87±3
E1	3	3.5±0.8	173±15
E1 (2.5,0)	3	8.1±0.8	84±7

[a] Software aperture in arcseconds.

[b] Position angle, north through east.

[c] Offset position : ($8\widehat{''}$ E(+),$6\widehat{''}$ N(+)).

blown by stellar winds (Forest and Shure 1986). In Figure 6.5(a) polarisation vectors are shown superimposed on the contour version of the K intensity image. The vectors over the extended emission to the west of Irs2 display a centro–symmetric pattern centred on Irs2. Since the average J–K and H–K colour–temperature of 1600K is too high for thermal emission from dust warmed by Irs2, we interpret all the J, H and K associated with Irs2, as reflection nebulosity. A $4\,\widehat{''}$ software aperture placed over the point source gives a linear polarisation of 10% and a position angle parallel to that in the reflection nebula $6\,\widehat{''}$W, where a $4\,\widehat{''}$ aperture yields a linear polarisation of 13%. Figure 6.6(a) is an east–west H–K colour profile through Irs2 and its reflection nebula, showing that the nebula becomes bluer towards the west. Our treatment of these observations will follow that of Castelaz *et al.* (1985 and 1986), who applied Hubble's relation (Hubble 1922) to the infrared reflection nebula surrounding GSS30 in Ophuichus and SGS1 in NGC1333.

Geometry of Irs2

Applying Hubble's relation to the J,H and K images of the reflection nebula, provides a means to deredden the illuminating source and thereby infer its intrinsic properties. The measured flux in a reflection nebula at a distance r from an illuminating source with flux $f_\star$ is $f_\star/4\pi r^2$, if the scatterers in the nebula scatter isotropically and have an albedo of one. Using the measured surface brightness in the nebula, Hubble's relation places a lower limit on $f_\star$. The faintest apparent magnitude of the source illuminating the reflection nebula is given by:

$$m_\star(\lambda) = m_n(\lambda) - 5\log_{10}(\theta) + k \tag{6.7}$$

where $m_\star(\lambda)$ is the apparent magnitude of the source in the direction of the reflection nebula, θ (in arcseconds) is the angular distance to the measured point in the nebula with surface brightness $m_n(\lambda)$, and $k = -2.75$ for isotropic single scattering from a unit optical depth nebula. The apparent magnitude of the source calculated from Hubble's relation provides a lower limit to the actual apparent 'dereddened' magnitude of the source required to illuminate the nebula to an angular distance θ. Table 6.5 summarises the results of applying Hubble's relation to Irs2 at J,H and K. The observed J, H and K magnitudes of Irs2 are fainter than those calculated from Hubble's relation, implying the presence of optically–thick material along the line–of–sight to the illuminating

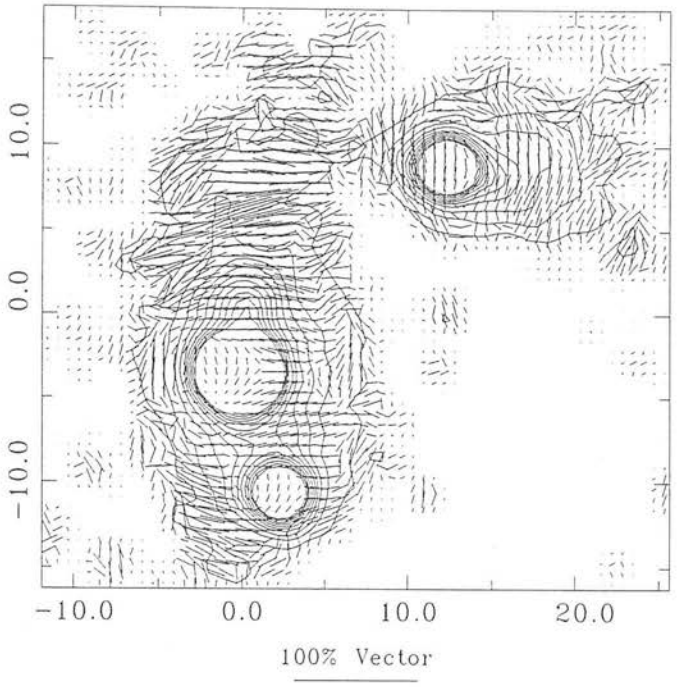


Figure 6.5(a) : High–resolution (0.6 $\widehat{''}$ pixel^{-1}) K (2.2 μm) imaging polarimetry of OMC 2 Irs2/4. Offsets are in arcseconds from Irs4 N.

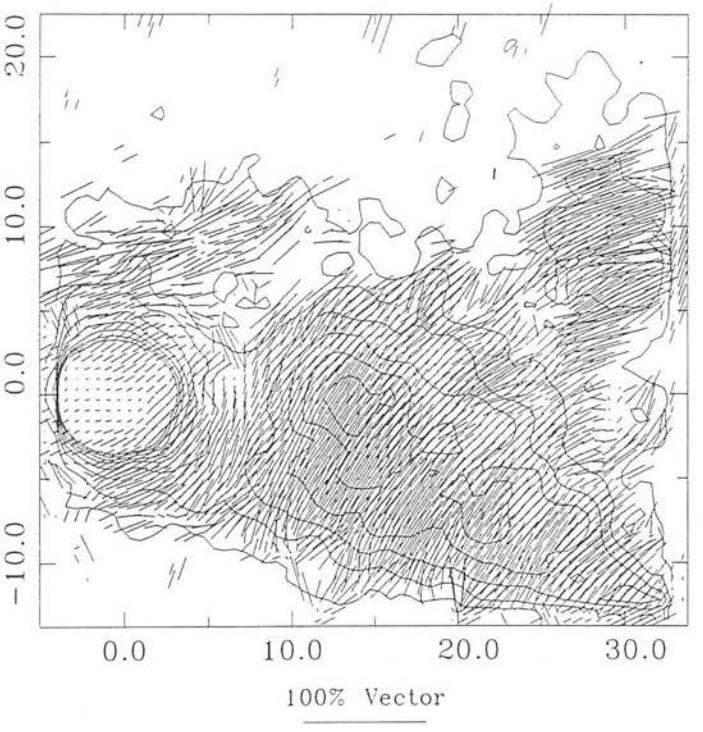


Figure 6.5(b) : High–resolution (0.6 $\widehat{''}$ pixel^{-1}) K (2.2 μm) imaging polarimetry of the OMC 2 Irs1/3 region. Offsets are in arcseconds from Irs3.

The polarisation electric vectors are shown superimposed on the 2.2 μm contoured intensity images.

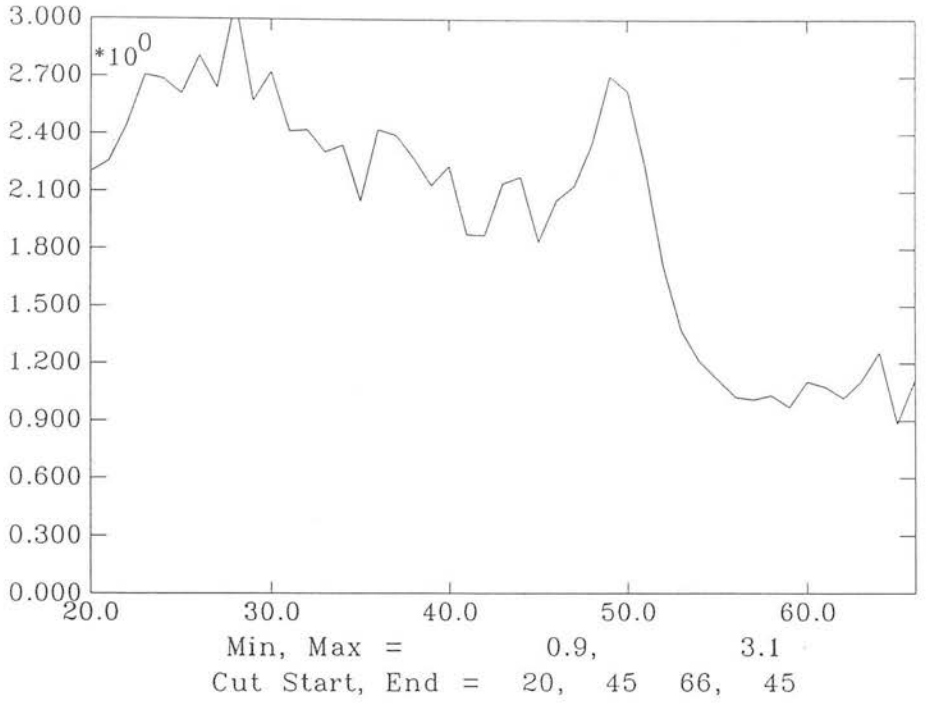


Figure 6.6(a) : H–K colour profile through OMC 2 Irs2 (east to west). The horizontal axis is in pixels ($0.625\,\widehat{''}$ pixel^{-1}), and is equivalent to a distance of $29\,\widehat{''}$.

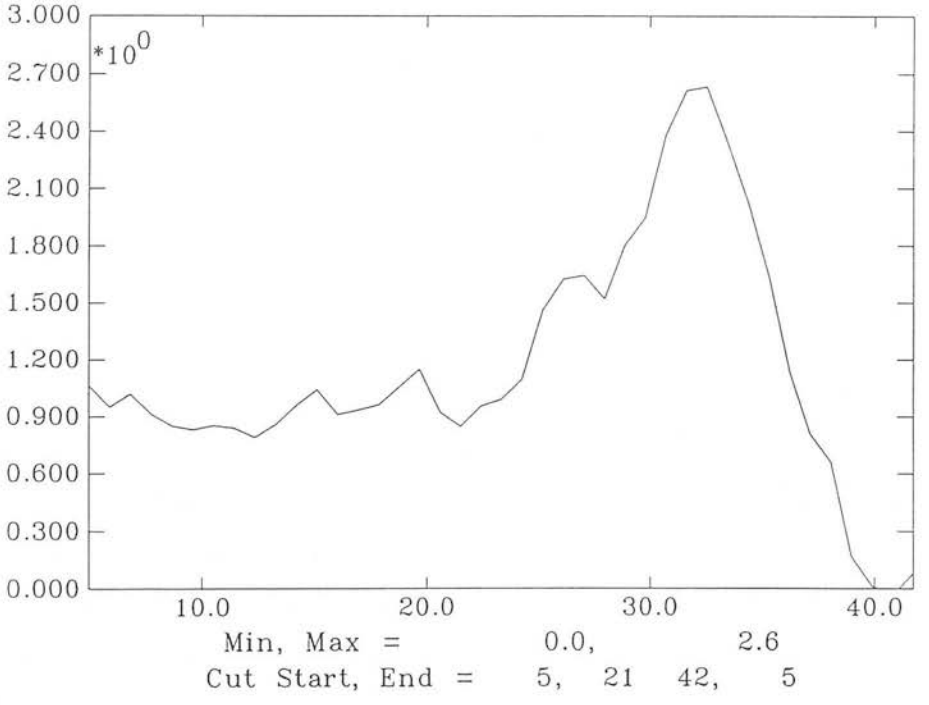


Figure 6.6(b) : H–K colour profile through OMC 2 Irs1 (north–east to south–west). The horizontal axis is in pixels ($0.625\,\widehat{''}$ pixel^{-1}), and is equivalent to a distance of $25\,\widehat{''}$.

Table 6.5: **Magnitudes from the Hubble relationship.**

Object	λ	θ^a	$M_n(\lambda)^b$	$M_*(\lambda)^c$	$M_o(\lambda)^d$	$A^e_{\lambda d}$	A^f_{Vd}
Irs2	J	10.5	17.8	9.9	>14.1	>4.1	>15
	H	11.0	16.2	8.3	12.2	3.9	22±2
	K	11.5	15.9	7.9	9.9	2.3	21±3
Irs4 N	H	13	18.9	10.6	14.3±0.5	3.8	21±3
	K	18	17.6	8.6	11.5±0.2	2.9	26±3
Irs1	J	30	18.2	8.0	>14.8	>6.8	>24
	H	34	17.2	6.8	12.1	5.3	30±3
	K	46	16.8	5.8	9.5	3.7	33±3

[a] Radial distance in arcseconds (±0.5) from the embedded source to $M_n(\lambda)$.

[b] 3σ limiting surface brightness limit in magnitudes (±0.2) per sq. arcsecond.

[c] Hubble magnitude (±0.2) of embedded source.

[d] Observed magnitude (±0.1) of embedded source into a 4″ aperture.

[e] $M_*(\lambda) - M_o(\lambda)$, the implied disk extinction at λ.

[f] The visual disk extinction,from the wavelength dependent extinctions of Rieke and Lebofsky (1985): $A_J=0.282A_V$, $A_H=0.175A_V$ and $A_K=0.112A_V$.

source, but not along the scattering path into the line–of–sight. Figure 6.6(a) shows an east–west profile through the H–K image of Irs2 and Irs4 at the declination of Irs2. From a maximum at Irs2 of H–K=2.7 the colour falls to H–K=1.0 at 5″W where it remains constant with increasing distance. As for SGS1 and GSS30 the structure of OMC 2 Irs2 is best explained by a source embedded in an optically–thick disk inclined to the line of sight, with its polar axis oriented east–west. In this scenario the increasing blueness of the nebula to the west is explained by the decreasing obscuring effect of the disk. The constant colour beyond 5″W implies that the western lobe of the reflection nebula extends beyond the disk and is unobscured. If the reflection nebula is bipolar then the absence of the eastern lobe from the images would require the disk to be inclined to the line–of–sight.

We interpret the alignment of the polarisation vectors along the line–of–sight to the Irs2 point source (see Figure 6.5(a)) as due to preferential extinction by aligned elongated grains. Models of disk formation have collapse taking place preferentially along local magnetic field lines creating a disk structure oriented perpendicular to the initial field (Ward–Thompson *et al.* 1985, Mestel and Paris 1984, Mestel and Ray 1985). Due to rotational torques, the frozen field lines form a toroidal geometry. Possible magnetic alignment mechanisms include the Davis–Greenstein (Davis and Greenstein 1951) and Purcell spin–up (Purcell, 1975 and 1979) mechanisms. To obtain the observed position angles the long axis of the grains must be aligned perpendicularly to the magnetic field in the disk, which must therefore be in the same plane as the polarisation vectors. In this picture the scattering lobes lie along the polar axis of the disk and so the position angle of the polarisation vectors due to scattering along the polar axis would be be parallel to the vectors in the putative disk, as observed. The alignment mechanism must be very efficient or work for small fields to produce the 10% linear polarisation observed with a $4\widehat{''}$ software aperture. However, a linear polarisation of 21% has been observed with a $4\widehat{''}$ aperture over BN, which is also most probably due to dichroic absorption (Hough *et al.* 1986).

Nature of Irs2

Hubble's relation can be used to estimate the total extinction to Irs2. Magnitudes calculated from this relationship effectively remove the source from behind the disk, and so the difference of the observed magnitude (m_o) and Hubble magnitude ($m_\star$) gives the wavelength dependent extinction due to the disk. As the J–H, H–K and J–K Hubble colours are unaffected by the disk, they are due only to foreground cloud extinction and intrinsic source colour. Using the calculated Hubble magnitudes together with the interstellar extinction law of Rieke and Lebofsky (1985), the visual extinction due to the Irs2 disk is estimated to be A_V=21 (see Table 6.5). The Hubble colours of Irs2 are given in Table 6.7, as are the foreground extinctions (A_{Vf}) estimated from the calculated colours at J–H, H–K and J–K. The errors on A_{Vf} are standard deviations derived by using the intrinsic colours of main sequence stars from B0 to M0 (Kornneef 1983, see Table 6.6). Using all three colours the best estimate is A_{Vf}=10±4. (For a B0V A_{Vf}=12±4, while for an M0V star A_{Vf}=7±2.) An independent estimate of the total extinction along the line of sight to the source can be estimated from the

Table 6.6: **Main Sequence intrinsic colours.**

Star	J-K	H-K	K-L
B0	−0.17	−0.05	−0.03
A0	0.01	0.00	0.00
F0	0.20	0.04	0.02
K0	0.47	0.10	0.05
M0	0.83	0.18	0.13

observed H–K colour, by using the interstellar extinction law. Using the H–K colour is a compromise between the uncertainty in intrinsic colour (−0.05 for B0V to 0.18 for M0V), and possible contamination of the K–L colour by warm dust; K–L is less sensitive to intrinsic star colour. Table 6.8 summarises the various extinction estimates. The sum of the disk and foreground extinctions derived from the Hubble relation $A_{V(d+f)}$=31±4 is in good agreement with that estimated from the observed colour, $A_{V(H-K)}$=37±3.

By using the best estimate of extinction (A_V=34±3), the observed magnitudes of Irs2 are dereddened at J,H,K,L and M (measured by Pendleton *et al.*), and are consistent with observing a stellar photosphere up to filter K (see Table 6.9). Hot dust in the vicinity of Irs2 explains the brightening seen at L and M. The position of Irs2 in a two colour (H–K,K–L) diagram (Cohen and Kuhi 1979) is that of a main sequence photosphere overlaid by thermal emission from hot dust grains at about 800K (see Figure 6.7). The internally consistent extinction values result in a dereddened apparent magnitude of K=6.1±0.4. At the distance of OMC 2 (500 pc, Goudis 1982), this corresponds to an absolute magnitude of $M_k = -2.4 \pm 0.4$. For a stellar photosphere the luminosity (L), effective temperature (T_{eff}) and absolute magnitude are related by the expression:

$$L/F_\lambda(K) = 4\pi d^2 \sigma T_{eff}^4 / \pi B_\lambda(\mathrm{K}, T_{eff}) \qquad (6.8)$$

where F_λ(K) is the flux equivalent to the absolute magnitude at K, d the distance (10 parsecs) and B_λ(K,T_{eff}) is the value of the Planck function at 2.2μm for a star of effective temperature T_{eff}. Irs2 can be placed in a luminosity–effective temperature (or Hertzsprung–Russell) diagram by plotting the curve for $M_k = -2.4$. We estimate

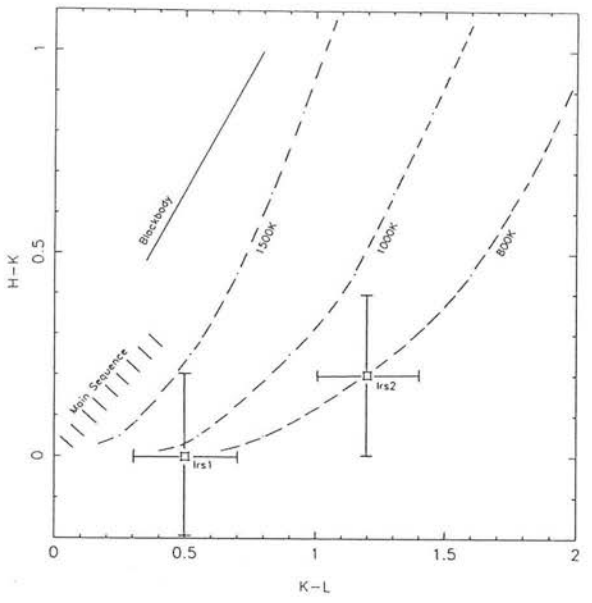


Figure 6.7 : Intrinsic two–colour (H–K, K–L) diagram for OMC 2 Irs1 and Irs2 (adapted from Cohen and Kuhi 1979). The hatched area represents the locus of main–sequence stars and the solid line represents the locus of blackbodies. The dashed loci refer to an A0 photosphere overlaid by thermal emission from hot dust grains at 800, 1000 and 1500K.

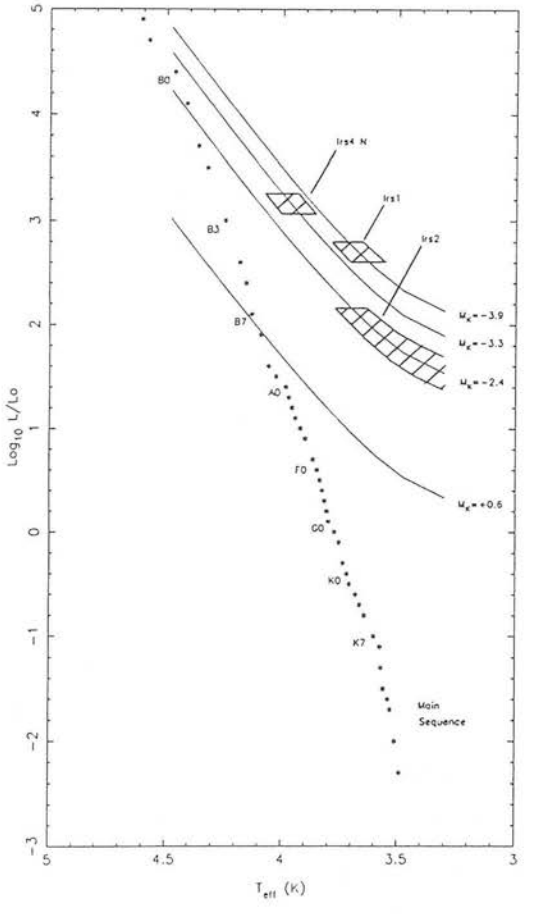


Figure 6.8 : Herztsprung–Russell diagram showing lines of luminosity as a function of absolute K magnitude (M_K). The positions of OMC 2 Irs1, Irs2 and Irs4 N are based on the intersection of their estimated far–IR luminosities with the absolute magnitudes estimated from the Hubble relationship. All three objects are placed above the main–sequence.

Table 6.7: **Hubble colours of the embedded sources.**

Object	*Colour*	$M_*(\lambda_1)$-$M_*(\lambda_2)$[a]	A_{Vf}
Irs2	(J-H)	1.6	13±3[b]
	(H-K)	0.5	6±2[c]
	(J-K)	2.0	10±2[d]
Irs4 N	(H-K)	2.0	32±3
Irs1	(J-H)	1.2	11±2
	(H-K)	1.0	16±2
	(J-K)	2.2	10±2

[a] Observed Hubble colour.

[b] The visual foreground extinction $A_{Vf} = 9.4$ [(J-H)-I(J-H)], where (J-H) is the observed Hubble colour, and I(J-H) is the assumed intrinsic colour of the source. The interstellar extinction law of Rieke and Lebofsky (1985) has been assumed.

[c] $A_{Vf} = 15.9$ [(H-K)-I(H-K)].

[d] $A_{Vf} = 5.9$ [(J-K)-I(J-K)].

an upper limit to the far–IR flux from Irs2 at about 150 $L_\odot$, since it was not detected in the 50 μm and 100 μm observations of Pendleton *et al.* This places Irs2 above the main sequence in the hatched area shown in Figure 6.8, in a region occupied by T Tauri and intermediate mass Herbig Ae/Be stars.

6.3.2 OMC 2 Irs4 Region

All the significant features in the region of Irs4 are seen in the K images (see Figure 6.3). Irs4 itself is resolved into two point sources, Irs4 S and Irs4 N, separated by about 4″. Some 8″ south of Irs4 S lies another point source which we will call E1 and which, together with Irs4 S, was imaged at 1 μm by Cohen and Frogel (1977). These 3 point sources lie in a north–south ridge of extended 2.2 μm emission. North–south intensity profiles through the low–resolution K image show that this ridge is spatially symmetric

Table 6.8: **Calculated visual extinctions.**

Object	A_{Vd}	A_{Vf}	$A^a_{V(d+f)}$	$A^b_{V(H-K)}$	A^c_V
Irs2	21±3	10±4	31±5	37±3	34±3
Irs4 N	24±4	32±3	56±5	56±5[d]	56±4
Irs1	32±3	12±3	44±4	41±3	43±3

[a] The sum of the disk and foreground extinctions.

[b] Total extinction to the embedded source calculated from the observed H-K colour and assuming an intrinsic colour of zero, $A_{V(H-K)}$=15.9E(H-K) from Rieke and Lebofsky (1985).

[c] Best estimate of the visual extinction to the embedded source.

[d] In this case the K-L colour has been used because of the uncertainty in H and since dereddening of the observed flux indicates that the photosphere is seen up to at least filter M. $A_{V(K-L)}$=18.5E(K-L) from Rieke and Lebofsky (1985).

about Irs4, extending about 20″ to the north and south, and with an east–west extent of about 7″. Photometry of the point sources and extended emission is detailed in Tables 6.2 and 6.3 respectively. (The nbL images are not sufficiently sensitive to detect the extended emission.) A plume–like feature appearing to originate at Irs4 N is the brightest part of a structure which arcs to the NW about Irs2. The morphology of this feature might be related to the Irs2 disk which possibly obscures the eastern lobe of an Irs2 bipolar reflection nebula. An inclined disk of the type proposed would obscure background nebulosity in a manner similar to that observed. It will be shown below that the nebulosity to the north of Irs4 probably lies behind Irs2. Other significant features seen in the 2.2 μm extended emission include diffuse emission extending 5″ south–east from Irs4 N and a circular nebula 4″ south of E1.

Geometry and Nature of Irs4

The orientation of the position angles and the high linear polarisation (20%–30%) of the vectors in the extended emission to the north and south of Irs4 (see Figure 6.5(a)) are indicative of a scattering mechanism, with either Irs4 S or Irs4 N as the illumination source. Futhermore, the orientation and length of the vectors along the line–of–sight

Table 6.9: **Dereddened magnitudes of the embedded sources.**

Object	λ	$M_o(\lambda)$	A_λ[a]	m[b]	M_k[c]
Irs2	J	>14.1	9.6±0.8	>4.5	
(A_V=34±3)	H	12.2	6.0±0.5	6.3	
	K	9.9	3.8±0.3	6.1	−2.4±0.4
	L	6.9	2.0±0.2	4.9	
	M[d]	5.0	0.8±0.1	4.2	
Irs4 N	H	14.3±0.5	9.8±0.7	4.5	
(A_V=56±4)	K	11.5±0.2	6.3±0.5	5.2	−3.3±0.5
	L	8.5	3.2±0.2	5.3	
	M	6.3	1.3±0.1	5.0	
Irs1	J	>14.8	12.1±0.9	>2.7	
(A_V=43±3)	H	12.1	7.5±0.5	4.6	
	K	9.5	4.9±0.3	4.6	−3.9±0.4
	L	6.7	2.6±0.2	4.1	
	M	4.7	1.0±0.1	3.7	
Irs4 S	J	13.2	7.2±0.9	6.0	
(A_V=26±3)	H	10.5	4.5±0.5	6.0	
	K	8.9	2.9±0.3	6.0	−2.5±0.3
	L	7.5	1.5±0.2	6.0	
	M	6.3	0.6±0.1	5.7	
E1	J	12.0	2.3±0.9	9.2	
(A_V=10±3)	H	10.7	1.4±0.5	9.0	
	K	10.2	0.9±0.3	9.1	0.6±0.3
	L	9.5	0.5±0.2	8.9	
Irs3	J	12.6	6.7±0.9	5.9	
(A_V=24±3)	H	9.4	4.2±0.5	5.2	
	K	7.7	2.7±0.3	5.0	−3.5±0.3
	L'[e]	4.6	1.3±0.2	3.3	
	M	3.2	0.6±0.1	2.7	

[a]Wavelength dependent extinctions from Rieke and Lebofsky (1985): A_J=0.282A_V, A_H=0.175A_V, A_K=0.112A_V, A_L=0.058A_V, and A_M=0.023A_V.

[b]Dereddened apparent magnitude.

[c]Derived absolute magnitude at K, $M_k = m - 5\log_{10}(\text{ dist. in pc.}) + 5$. The assumed distance to OMC 2 is 500pc.

[d]All the M magnitudes are from Pendleton *et al.* (1986).

[e]From Pendleton *et al.* (1986).

to Irs4 N are not inconsistent with scattering centred on Irs4 S. Similar observations at 3.8 μm but at much lower resolution, lead Pendleton *et al.* to suggest that Irs4 N is a reflection nebula illuminated by Irs4 S. Because of the consistency of position angle and polarisation across Irs4 N and out into the extended emission in the high–resolution 2.2 μm vector map, it seems most unlikely that Irs4 N is a background source viewed through a reflection nebula illuminated by Irs4 S. In fact the inconsistency of polarisation with position occurs at Irs4 S where a 4″ software aperture measures a linear polarisation of only 2.8% and a position angle of 0°. By comparison the same aperture measures 20%/87° and 30%/102° at Irs4 N and in the nebula to the north of Irs4 respectively (see Table 6.4). If we treat Irs4 S as a foreground source what remains is a bipolar reflection nebula illuminated by Irs4 N, with the high line–of–sight linear polarisation being due to dichroic absorption in a disk surrounding an embedded source, similar to Irs2. Since the bipolar reflection nebula is symmetrical about the illuminating source, we are probably seeing the disk edge–on and so a toroidal magnetic field in the disk with less obscured scattering lobes to the north and south would produce the observed vector pattern, as discussed previously in the case of Irs2. Although the 2.2 μm linear polarisation of 20% is very high for a dichroic mechanism, it is no higher than that observed at BN. If the polarisation of Irs4 N follows the same trend as BN, then we should expect to see the linear polarisation decrease to about 10% at 3.8 μm (Hough *et al.* 1988 and references therein). However Pendleton *et al.* observe values of 25% into a 4″ beam over Irs4 N at 3.8 μm. High–resolution imaging polarimetry longward of 3.5 μm is required to separate Irs4 N from the extended emission.

Support for this interpretation comes from attempts to construct a near–IR scattering model of Irs4 N. In this model Irs4 N is a reflection nebula with a diameter of ϕ″ illuminated by Irs4 S at a distance of d″. The solid angle subtended by the reflection nebula as seen from Irs4 S is then $\Omega = \pi(\phi/2)^2/d^2$. Assuming single isotropic scattering and unit albedo, the minimum difference in magnitude between the source and nebula is $\Delta m = 2.5\log_{10}(\Omega/4\pi)$. If it is assumed that the intrinsic H–K colour of Irs4 S is zero and that the observed H–K colour is due solely to interstellar extinction, the extinction law of Rieke and Lebofsky (1985) gives J,H,K,L and M dereddened apparent magnitudes of about 6.0, consistent with observing a stellar photosphere up to at least filter M (see Table 6.9). By assuming that d is equal to the projected distance of 4″ and that

Table 6.10: **Irs4 N Near-IR Scattering Model.**

λ	m_o[a]	m_{neb}[b]	
		A_V=0	A_V=40
H	14.0	10.5	17.5
K	11.3	10.5	15.0
L	8.5	10.5	12.8

[a] Magnitude observed in a 4$\widehat{''}$ software aperture.

[b] Modelled magnitude of the Irs4 N reflection nebula.

ϕ=2$\widehat{''}$, the FWHM of Irs4 N, then $\Delta m = -4.5$. Consequently the *maximum possible* apparent magnitude of Irs4 N at H,K and L is 10.5 (see Table 6.10). Even without any intervening extinction the observed magnitude at L is already 2 magnitudes brighter than predicted by the scattering model. Although the observed H and K magnitudes are in themselves consistent with the model, the observed H–K and K–L colours are far too red. When the 40 magnitudes of visual extinction required to reproduce the observed colours are put into the model (either in the form of foreground extinction, extinction between the illuminating source and nebula, or combination of both), the brightness of the reflection nebula becomes completely untenable. In summary, a simple but very conservative near–IR scattering model does not work because Irs4 N is too bright and too red. Perhaps even more convincing evidence that the near–IR flux from Irs4 N is intrinsic comes from the observation that it is a point source at K and L.

An H–K profile through Irs4 N and its reflection nebula to the north peaks at H–K=2.8 at Irs4 N and then falls to a constant value of H–K=1.8 at about 15$\widehat{''}$north of Irs4 N, in a similar way, but less steeply than the H–K profile through Irs2 and into its western reflection nebula. Because the obscuring disk is viewed edge–on it is not surprising that the H–K colour gradient is less steep than in the case of Irs2, where the evidence suggests that the disk is tilted such that its obscuring effect on the western scattering lobe is less. Given the evidence that Irs4 N is the illuminating source for the Irs4 nebula, Tables 6.5,6.7,6.8 and 6.9 list the results of applying Hubble's relationship

to Irs4 N and the northern reflection nebula. Calculated disk (A_{Vd}=24±4) and foreground (A_{Vf}=32±3) extinctions imply a total extinction to Irs4 N of $A_{V(d+f)}$=56±5, in excellent agreement with that derived from the observed K–L colour of the source. Dereddening the observed magnitudes results in H,K,L and M magnitudes of about 5.2, again consistent with observing a stellar photosphere up to at least filter M. When the resulting absolute magnitude of $M_k = -3.3\pm0.5$ is plotted in the H–R diagram, the intersection of the measured far–IR luminosity of 1500 $L_\odot$ (Pendleton *et al.*) with locus of constant M_k, places Irs4 N above the main sequence as shown in Figure 6.8. The dereddened apparent magnitudes show no evidence of any IR excess due to hot dust. The lack of significant continuum emission at 10 μm (see Figure 6.9) could indicate that the source is embedded inside a dust free cavity.

Assuming Irs4 S to have an intrinsic H–K colour of zero, the dereddened apparent magnitudes at J,H,K,L and M, of about 6.0, are consistent with observing a stellar photosphere up to filter M (see Table 6.9). Although the tongue of illumination seen extending 5$\widehat{''}$ south from Irs4 S has a linear polarisation of about 20%, the position angles could result from illumination by either Irs4 N or S, however the H–K colour of about 1.3 is much bluer than the extended emission thought to be illuminated by Irs4 N. Any disk–like structure around Irs4 S could well obscure Irs4 N. The mid– to far–IR photometry of Irs4 S by Pendleton *et al.* indicates that it is far less luminous than Irs4 N. Assuming an upper limit of 150 $L_\odot$, the dereddened absolute magnitude of $M_k = -2.5\pm0.4$ places Irs4 S on the H–R diagram (Figure 6.8) in the same position as Irs2, and so it too is perhaps a pre–main sequence T Tauri or Herbig Ae/Be star. Although we have found no evidence to suggest a connection between Irs4 N and S, the 4$\widehat{''}$ separation of the two objects gives a projected distance of 2,000 a. u. at 500 pc. and so Irs4 could possibly be a pre–main sequence binary. A further 8$\widehat{''}$ south of Irs4 S lies the point source E1, which has an H–K colour second only to the known near–IR sources. At H–K=0.5, E1 is probably too red to be a foreground star and not red enough to be a background star seen through the cloud, which has an estimated visual extinction of 30–90 magnitudes (Westbrook *et al.* 1976,Thronson *et al.* 1978), equivalent to an extrinsic colour of H–K=2–6. Of course if the cloud is clumpy, E1 could be seen through a gap, but its proximity to Irs4 is possibly significant. It is interesting to note that the H–K colour of the sources decreases approximately linearly with distance south from Irs4 N, through Irs4 S, to E1. It is tempting to suggest that

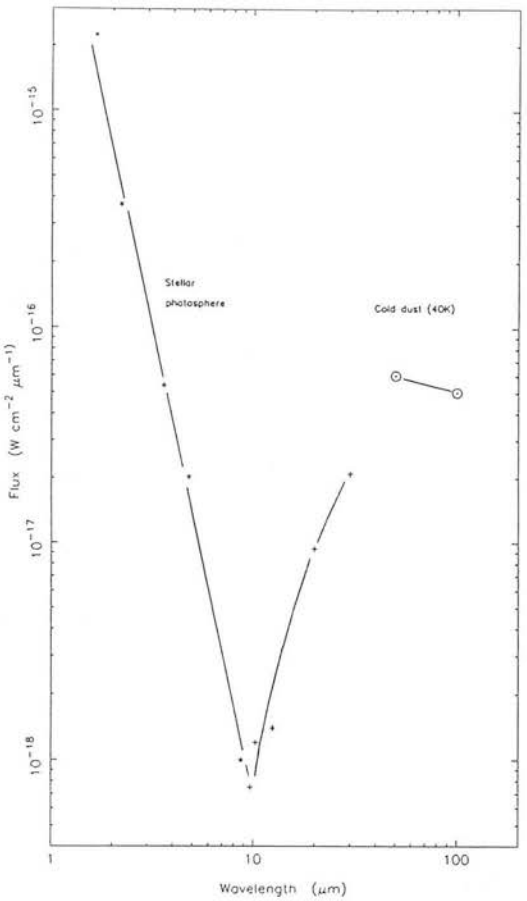


Figure 6.9 : The intrinsic flux from OMC 2 Irs4 N based on a visual extinction of $A_V = 56$. A $4\widehat{''}$ beam was used from 1–20 μm and a $30\widehat{''}$ beam at 50 μm and 100 μm. The 4–100 μm observations are from Pendleton *et al.* (1986).

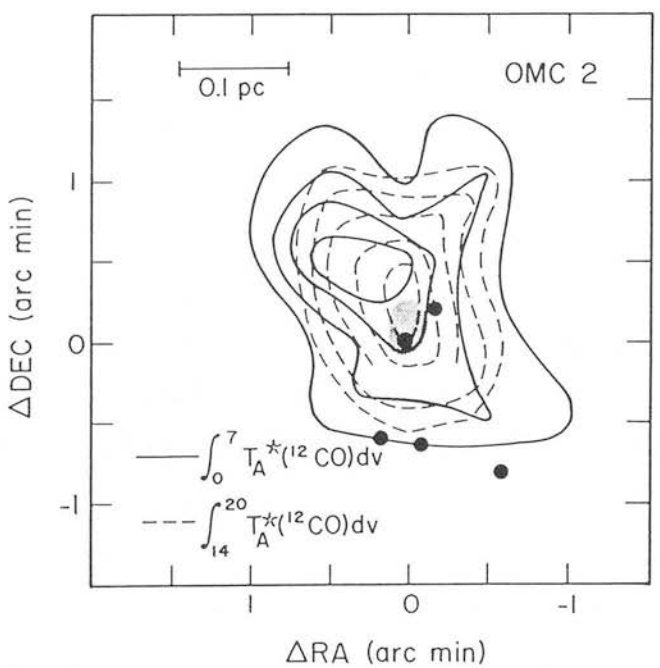


Figure 6.10 : (Fischer *et al.* 1985). Map of the integrated intensity of the ^{12}CO emission of OMC 2 in the blueshifted (solid contours) and redshifted (broken contours) wings. The (0,0) position of the map is at OMC 2 Irs4 N. The shaded region marks the approximate extent of the H_2 S(1) line–emission (Fischer *et al.* 1980). The positions of Irs1, 2, 3, 4 and 5 are marked by filled circles.

E1 is perhaps the least embedded member of an Irs4 triple system. Dereddening E1 on the basis of an H–K colour of about zero gives an absolute magnitude of M_k=0.6±0.3 (see Table 6.9), if E1 is in OMC 2. Its location in the H–R diagram is then that of a B8 main sequence star or lower luminosity pre–main sequence star. Both Irs4 S and E1 have small linear polarisations (~3%) compared to Irs2 (10%) and Irs4 N (20%), where we have suggested that disk structures are implicated.

Evidence of Young Stellar Object Activity

Several of the classic signs of YSOs are present in the Irs4 region: molecular H_2 emission, a bipolar CO high velocity outflow and an H_2O maser. Figure 6.10 is taken from Fischer *et al.* (1985) and shows the blue– and redshifted wings of the CO outflow. The shaded region marks the extent of the molecular H_2 S(1) v=1–0 line emission region (Fischer *et al.* 1980) the discovery of which prompted the search for an excitation mechanism and which lead to the discovery of the CO outflow. Observations with 15″ and 34″ beams demonstrated that the S(1) was extended on a scale of 15–30″, weighted to the north of Irs4. UKIRT service observations with 5″ beams (Burton, priv. comm.), confirmed that the emission is extended both to the north and south of Irs4 N. From moderate spectral resolution observations in the near–IR Thronson and Thompson (1982) identified at least 8 members of the v=1–0 and v=2–1 rotation series of H_2. They argue that the line ratios and intensities are due to a shock excitation mechanism, however their spectrum is unusual and the signal–to–noise is low, and so this interpretation is open to doubt. By using pairs of lines formed in the same upper level but emitted at different wavelengths, Thronson and Thompson estimate the amount of extinction to the H_2 source to be A_V=28±6, assuming an interstellar extinction law of the form A_λ=0.4$A_V\lambda^{-1.9}$ (μm). From their quoted line fluxes we have re–estimated the amount of extinction using the law of Rieke and Lebofsky (1985), which can be put in the form A_λ=0.44$A_V\lambda^{-1.64}$ between 1 and 4 μm. A weighted average of the 3 line pairs (S(1)/Q(3), S(2)/Q(4) and S(3)/Q(5)) gives A_V=24±5 (although the individual values are respectively,with decreasing confidence, 16, 35 and 50). This value agrees within the errors with the foreground extinction to Irs4 N of $A_{Vf} = 32 \pm 3$, derived from Hubble's method (see Table 6.8). Given the line–of–sight and optical depth coincidence, together with the similar spatial morphologies of the S(1) and 2.2 μm extended emission, we conclude that the two emissions are associated

(although the S(1) line emission is not bright enough to contribute significantly to the 2.2 μm continuum).

An approximate north–south projected outflow axis is indicated by a map of the integrated intensity in the ^{12}CO emission of OMC 2 in the blue– and redshifted emission wings (Figure 6.10, from Fischer *et al.* 1985). The positional uncertainty of the CO map relative to the near–IR sources is $\pm 7''$ and on the basis of the map alone it is not possible to choose between Irs4 N, Irs4 S and Irs2 as a potential driving source. However, given that Irs4 N illuminates a north–south oriented bipolar reflection nebula, that it is probably associated with molecular H_2 emission and that it is the most luminous source in OMC 2, we conclude that it also drives the outflow. If the outflow is radiatively driven by a central source of luminosity $L_\star$, the momentum transfer $\dot{p}$ from the radiation to the surrounding matter is constrained by:

$$\dot{p} \leq \frac{L_\star}{c}\tau_R \tag{6.9}$$

where $L_\star/c$ is the force that could be exerted if every photon emitted by the source were absorbed or scattered once by the material in the flow, and τ_R is the effective optical depth which couples the radiation to the surrounding material. Using the luminosity $L_\star$=3x10^4 L$_\odot$ for OMC 2 from Smith *et al.* (1979), Fischer *et al.* (1985) deduce a value of τ_R=1, and conclude that radiation pressure could accelerate the outflow. However the more appropriate Irs4 N luminosity of 1500 L$_\odot$ gives $\tau_R \geq$23. Most of the scattering is expected to occur at near–IR wavelengths (Solomon *et al.* 1981). We measure a foreground extinction to the northern lobe of the bipolar outflow of A_{Vf}=32 which is equivalent to a 2.2 μm optical depth of $\tau \sim A_k \simeq 3$ and so the optical depth in the lobe must be τ <3. Consequently it seems unlikely that radiation pressure can drive the OMC 2 outflow.

An admittedly speculative but nevertheless self–consistent model of Irs4 has Irs4 N as the central source driving a high–velocity CO outflow. Irs4 N is a pre–main sequence star, possibly a Herbig Ae/Be object, embedded in a protostellar disk oriented approximately east–west, and illuminates a north–south bipolar reflection nebula. It is perhaps most satisfying to explain the molecular H_2 emission in terms of emission from cooling gas that has been shocked by the outflow from Irs4 N, where the observed north–south extension arises naturally from the east–west orientation of the collimating disk. A stellar wind could also clear a dust–free cavity around the source, hence

explaining the lack of 10 μm continuum.

6.3.3 OMC 2 Irs1 Region

OMC 2 Irs1 is the 2.2 μm point source lying about one arcminute south–east of Irs4. The low–resolution K image (see Figure 6.1) shows that Irs1 is surrounded by extensive and spatially complex 2.2 μm continuum emission. Although the conical nebula to the north–east of Irs1 is by far the brightest 2.2 μm continuum emission, the nebulosities to the north and in a finger–like protrusion to the north–west, have similar extents, reaching the 3σ surface brightness limit at about 40$\widehat{''}$ from Irs1. Fainter nebulosity is also apparent to the south of Irs1, between the stars π2007 and π1988. However it is noticeably absent from the wedge shaped regions to the east and west of Irs1 (see Figure 6.11). Even though the quality of the low–resolution H image is poor by comparison with the K image, due to an interruption in data–taking, the morphology of the two images is similar. At J however, only the extended emission to the north–east and south is visible, while Irs1 itself is not detected. Photometry of the extended emission at J,H and K is summarised in Table 6.3. Low–resolution imaging polarimetry confirms that much of the 2.2 μm continuum within 40$\widehat{''}$ of Irs1 is due to illumination by Irs1. Polarimetry with 12.6$\widehat{''}$ beams indicates that the nebulosity to the south and south–west of Irs1, which the IRCAM imaging polarimetry did not cover, is also due to scattering centered on Irs1. The high–resolution 2–image K mosaic of Irs1 and the bright conical shaped reflection nebula to the north–east is shown in Figure 6.4. Irs3 is the bright point source to the west. Irs1 is "teardrop" shaped, with its axis aligned with the bright reflection nebula. Apart from the 'neck' in the nebula about 6$\widehat{''}$ north–east of Irs1, the surface brightness is approximately constant at K=14.0 per sq. arcsecond (40σ) out to 25$\widehat{''}$, falling to within 6σ of the sky at 32$\widehat{''}$, and then more slowly to within 3σ of the sky brightness at 46$\widehat{''}$. A single high–resolution H image displays a morphology similar to the corresponding K image.

Geometry of Irs1

Figure 6.6(b) shows a profile through the high–resolution H–K image along the (north–east) axis of the bright reflection nebula. From a maximum of H–K=2.7 at Irs1, the

colour falls to H–K=1.0 at 7″, where it remains constant out to the edge of the image at 18″. A similar profile through the low–resolution H–K image continues this trend out to 40″. Significantly the colour shows a discontinuity at about 5″ north–east of Irs1, where the nebula apparently 'detaches' from Irs1 (see Figure 6.4). At this point the colour increases to H–K=1.6. A systematic increase is also observed in the H–K colour from 1.0 over the bright nebula to about 1.6 in the fainter nebulosity 25″ to the north of Irs1. We interpret the decrease in colour to the north–east along the bright reflection nebula in the same manner as the colour gradients seen in Irs2 and Irs4, namely decreasing obscuration by a circumstellar disk. The increase in colour at the point where the north–east nebula detaches from Irs1 probably indicates additional obscuration by a ridge of material which also separates the north and north–western nebulosities. If the northern nebula is dereddened relative to the north–eastern nebula by assuming that the H–K colour differential of 0.6 is due simply to additional line of sight obscuration, it is brightened by about one magnitude at K, and attains a surface brightness only slightly greater than the north–eastern nebula. Similarly the north–western nebula also becomes one magnitude brighter. High–resolution imaging polarimetry of this region (see Figure 6.5(b)) shows a centro–symmetric vector pattern consistent with illumination by Irs1. Although knots of linear polarisation of up to 50% occur, we find that the polarisation is approximately constant at 30% along the axis of the of the north–eastern nebula. The linear polarisation also averages 30% in the northern lobe. These high values probably indicate large scattering angles, implying that the axes of the nebulae lie close to the plane of the sky.

The extensive but faint nebulosity seen to the south is also illuminated by Irs1. A contoured K intensity map of the region (see Figure 6.11) is best understood in terms of a bipolar reflection nebula, with the illumination axis aligned approximately north–south. The absence of extended emission in the wedge–shaped regions to the east and west of Irs1 is then naturally explained as the result of an east–west aligned circumstellar disk in which Irs1 is embedded. Given the suggested bipolar geometry, we might expect the southern scattering lobe to have the same unobscured brightness as the northern lobe. However, the colours of the southern lobe are similar to those of the bright north–eastern reflection nebula (see Table 6.3), and so its faintness is unlikely to be due to differential extinction by the disk or line of sight material. Possible explanations include, absence of scatterers in the southern lobe or perhaps clumps of obscuring

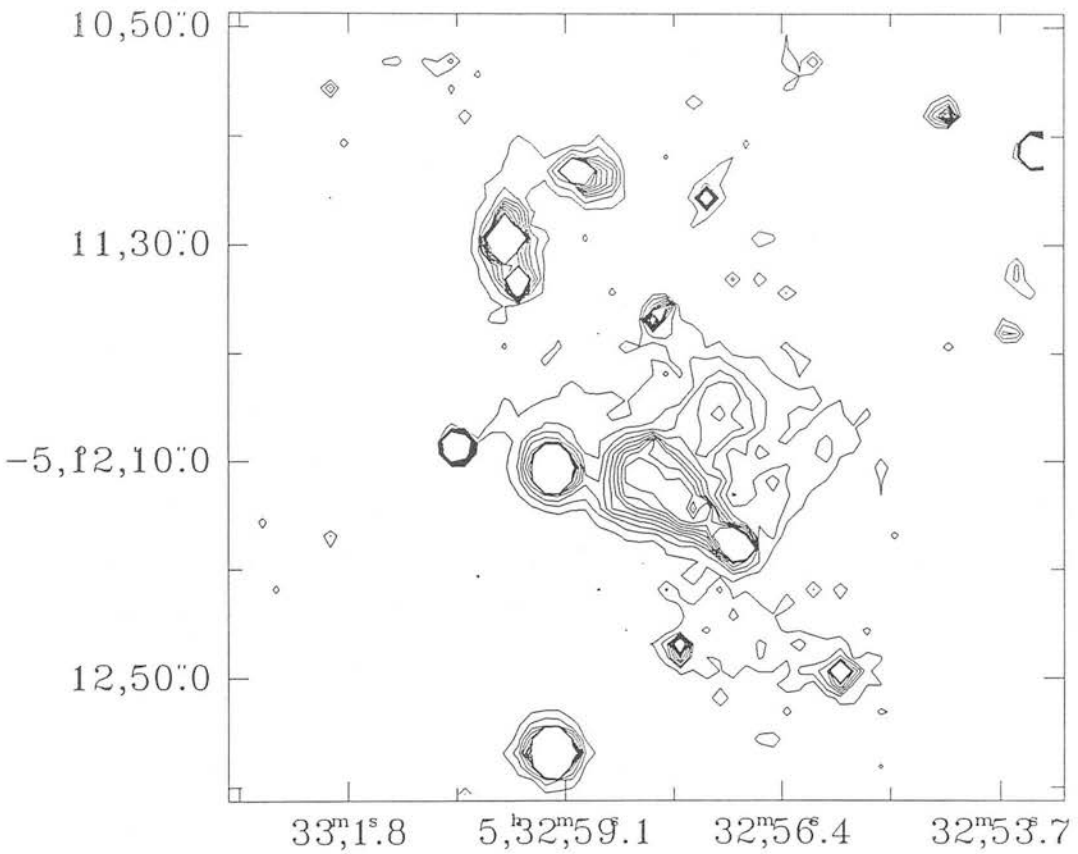


Figure 6.11 : A contour plot of the low–resolution (2.5 $\widehat{''}$ pixel^{-1}) K (2.2 μm) image of OMC 2. The lowest contour level is 16.2 magnitudes per sq. arcsecond, followed by 15.6, 15.2, 14.9, 14.7, 14.5, 14.3, 14.2 and 14.1.

matter close to Irs1; these clumps must not introduce selective extinction. Because of the evidence suggesting non–disk related obscuration in the other directions, Hubble's relationship is applied only to the north–eastern reflection nebula (see Tables 6.5,6.7,6.8 and 6.9). Calculated disk (A_{Vd}=32±3) and foreground (A_{Vf}=12±3) extinctions imply a total visual extinction to Irs1 of $A_{V(d+f)}$=44±=4, in good agreement with that derived from the observed H–K colour of the source. Making the simplifying assumption of circular symmetry in the Irs1 H–K colour profile, the weighted visual extinction into a 30$''$ beam centred on Irs1 becomes about A_V=13. This value is in good agreement with A_V=15 derived by Pendleton *et al.* from the far–IR optical depth averaged over a 30$''$ field of view containing Irs1.

Nature of Irs1

Dereddening the observed magnitudes of Irs1 at J,H,K,L and M, using the best estimate of visual extinction to the source (A_V=43±3), gives results which are consistent with observing a stellar photosphere up to filter K (see Table 6.9). The L and M magnitudes imply the presence of hot dust at an L–M colour temperature of 800±100K. By modelling the 3.8 μm continuum scattered from the north–eastern nebula, Pendleton *et al.* find that, to the extent it can be described by a single blackbody temperature, the illumination source is best characterised by an object at ~1000K. Given the different nature of our two methods, such agreement is gratifying. The position of Irs1 in a two colour (H–K,K–L) diagram is that of a main sequence photosphere overlaid by thermal emission from hot dust at about 800K (see Figure 6.7). At the distance of OMC 2 the dereddened apparent magnitude of K=4.6±0.4, corresponds to an absolute magnitude of $M_k = -3.9 \pm 0.4$. This places Irs1 above the main sequence in the H–R diagram (see Figure 6.8), in a region occupied by Herbig Ae/Be stars.

6.3.4 OMC 2 Irs3 Region

Our high–resolution 2.2 μm imaging polarimetry of OMC 2 Irs3 (see Figure 6.5(b)) shows a centro–symmetric polarisation vector pattern clearly centred on Irs3. About 8$''$ south of Irs3 the linear polarisation approaches 40–50%, typical of large angle scattering. Aperture polarimetry on source (see Table 6.4) yields a linear polarisation of

3.6±0.2% with a position angle of 105±5° in agreement with previous determinations, (Dyck and Lonsdale, 1979). Mid– to far–IR photometry of Irs3 by Pendleton *et al.* shows that it is not a major luminosity source and has a 1–20 μm integrated luminosity of 50 $L_{\odot}$. If Irs3 is dereddened by assuming an intrinsic H–K colour of zero, its position on the H–R diagram gives T_{eff} <1000K, inconsistent with the assumed intrinsic colour. We should not be surprised at this failure since Irs3 has a 2–20 μm flux distribution consistent with a single blackbody energy distribution with a colour temperature of ~500K (Wilner *et al.* 1982).

6.4 Summary

The wide spatial coverage and high spatial resolution, together with the imaging polarimetry capabilities offered by IRCAM, has proved crucial in complementing extensive observations at other wavelengths and allowing us to disentangle the complex region of near–IR embedded sources and reflection nebulae, known as OMC 2.

Using Hubble's relationship we find that the objects Irs1, Irs2 and Irs4 N are best explained as sources embedded in circumstellar disks and illuminating near–IR reflection nebulae. Due to the orientation of the disk, only the western scattering lobe of Irs2 is seen. Since both the north and south lobes of Irs1 and Irs4 N are visible, we are probably viewing the putative disks more edge–on. We find no evidence for alignment of the disk structures with the local magnetic field threading OMC 2, for which Dyck and Lonsdale (1979) give a position angle of ~120°. Dichroic absorption in the disks of Irs2 and Irs4 N seems the most reasonable explanation of the high 2.2 μm linear polarisations (10% and 20% respectively) seen along the line–of–sight to the compact sources. If grain alignment is caused by a toroidal magnetic field in a circumstellar disk seen edge–on, then the dichroic polarisation vectors would be produced parallel to the disk, and parallel to the scattering vectors in the less obscured reflection lobes along the polar axis of the disk, as observed.

Irs1, Irs2 and Irs4 N are dereddened using the self–consistent visual extinctions found from Hubble's relationship and the observed near–IR colours. On the H–R diagram Irs1 lies above the main sequence in a position occupied by intermediate Herbig Ae/Be objects, and its dereddened near–IR colours are consistent with a stellar

photosphere overlaid by thermal emission from hot dust grains at ~1000K. Similarly, Irs2 lies in a region of the H–R diagram occupied by Herbig Ae/Be or possibly T Tauri–type objects, its colours too are those of a stellar photosphere overlaid by hot dust at ~800K. Irs4 N is the most luminous source in OMC 2 and its near–IR flux is intrinsic and not scattered from Irs4 S as previously suggested. It lies above the main sequence in a region occupied by Herbig Ae/Be objects, although its dereddened colours show no sign of hot dust, in contrast to Irs1 and 2. The northern scattering lobe is associated with molecular H_2 emission by virtue of a similar morphology and line–of–sight optical depth. This emission is probably due to shock–excitation in the high–velocity CO outflow most likely driven by Irs4 N.

Appendix A

Thermal Radiation

Thermal radiation is emitted by a body by virtue of its temperature. The term emissive power is used to denote the emitted thermal radiation leaving a surface per unit time per unit surface area. The total hemispherical emissive power of a surface is all the thermally emitted energy, summed over all directions and wavelengths, and is usually denoted by E. The total emissive power is dependent on the temperature of the emitting surface, the substance of which it is composed, and the nature of the surface structure. In general, the emissive power of a given surface element is directionally dependent. The radiation intensity I describes the directional distribution of radiant energy leaving a surface (either by emission or reflection). With reference to Figure A.1, the intensity at a point P at a distance r and in a direction θ, due to radiation originating from a point in a surface, is defined as :

$$I = \frac{df}{d\omega}\frac{1}{\cos\theta} = \frac{df}{da}\frac{r^2}{\cos\theta} \tag{A.1}$$

The quantity $df/d\omega$ represents the radiated flux per unit solid angle, while df/da is the areal flux density at P of the flux emanating from Q. Since all the flux leaving the surface at Q must pass into the hemisphere above it, then :

$$f = \int I\cos\theta d\omega = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2} I(\theta,\phi)\sin\theta\cos\theta d\theta d\phi \quad \text{where } d\omega = \sin\theta d\phi d\theta \tag{A.2}$$

Radiation is termed diffuse if I is uniform (independent of θ and ϕ). Thus, for diffuse radiation :

$$f = I\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2}\sin\theta\cos\theta d\theta d\phi = I\pi \tag{A.3}$$

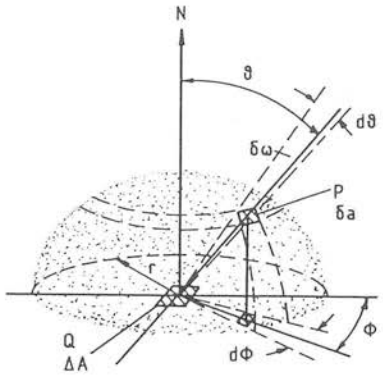


Figure A.1

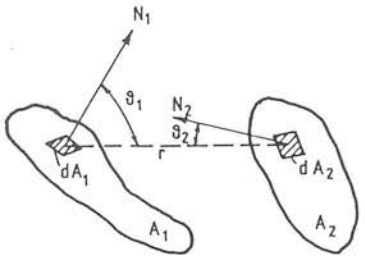


Figure A.2

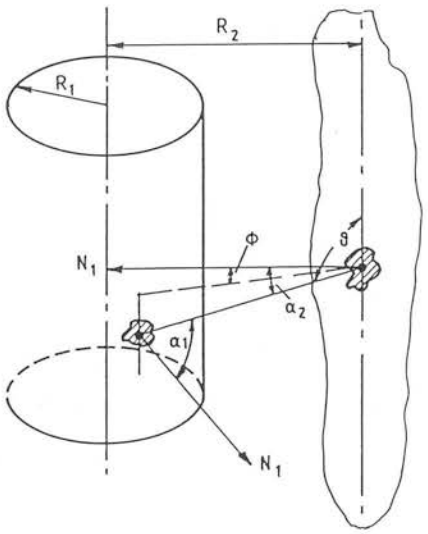


Figure A.3

Also, Equation A.1 shows that if I is constant the flux intercepted by an area element varies as $\cos\theta$ and inversely as r^2. Such a radiator is known as a Lambertian radiator. In the case where the radiant flux f is due solely to thermal emission from a surface, the flux is the hemispherical emissive power E, and the associated intensity is symbolised by $I_e(\theta,\phi)$. A surface is termed a diffuse emitter if $I_e(\theta,\phi)$. Then :

$$E = \pi I_e \tag{A.4}$$

Similar relationships exist for all the energy leaving (emitted and reflected) leaving a surface. When radiation is incident upon a surface, part of it may be reflected from the surface, part of it may be absorbed by the surface, and part may be transmitted through the surface. Applying conservation of energy gives :

$$\rho(\lambda) + \alpha(\lambda) + \tau(\lambda) = 1 \tag{A.5}$$

where $\rho(\lambda)$ is the monochromatic reflectivity, the fraction reflected at λ
$\alpha(\lambda)$ is the monochromatic absorbtivity, the fraction absorbed at λ
$\tau(\lambda)$ is the monochromatic transmittivity, the fraction transmitted at λ

The total surface properties (ρ,α and τ) are the monochromatic properties summed over all wavelengths. In general the monochromatic properties depend upon the surface composition, the nature of the surface (roughness etc.), and its temperature.

The fundamental equation for heat transfer by radiation from a black body (a perfect radiator) is the Stefan–Boltzmann law :

$$E_b = \int_0^\infty E_{b\lambda}(\lambda,T)d\lambda = \sigma T^4 \tag{A.6}$$

$$\text{where} \quad E_{b\lambda}(\lambda,T) = \frac{2\pi hc}{\lambda^5}\frac{d\lambda}{exp(hc/\lambda kT)-1} \tag{A.7}$$

and where E_b is the blackbody total emissive power
$E_{b\lambda}$ is the Planck radiation function (monochromatic E_b)
T is the absolute temperature of the emitter
λ is the wavelength
c is the velocity of light
σ is the Stefan–Boltzmann constant ($5.67\text{x}10^{-8}\text{Wm}^{-2}\text{K}^{-4}$)
h,k are the Planck and Boltzmann constants

A perfect blackbody is defined as one which absorbs all incident radiation, ie. $\alpha(\lambda) = \alpha = 1$ or $\rho(\lambda) = \rho = 0$. As may be deduced from the Planck radiation function (Equation A.7), the monochromatic blackbody emissive power exhibits a peak value that increases with T and which occurs at shorter wavelengths as T increases. This is known as Wien's Displacement Law :

$$\lambda_{max}T = 2898 \quad \mu\text{m K} \tag{A.8}$$

It follows that at low temperatures the radiation is almost entirely confined to longer infrared wavelengths. Most bodies are not perfect radiators and emit radiation at a lower rate than a blackbody thus :

$$E = \sigma\varepsilon T^4 \tag{A.9}$$

where the emissivity ε is defined by :

$$\varepsilon(\lambda,T) = \frac{E_\lambda(\lambda,T)}{E_{b\lambda}(\lambda,T)} \tag{A.10}$$

A special type of non-black surface, called the Ideal Gray Body, is defined as one for which the monochromatic emissivity is independent of wavelength, ie. the ratio of E_λ to $E_{b\lambda}$ is the same for all wavelengths at a given temperature. The properties of emissivity and absorbtivity are related by Kirchoff's law (see for example Chapman 1984) which states :

$$\alpha(T) = \varepsilon(T) \tag{A.11}$$

when either of the following conditions is met :

1. If the radiation incident upon the receiving surface (known as the irradiation) has the same spectral distribution as that of the emission from a blackbody at the same temperature.

2. If the receiving surface is an ideal gray body.

The first condition restricts the irradiation source to a blackbody at the same temperature as the receiver and therefore finds no application in heat transfer since thermal equilibrium rarely exists. The second condition does not require thermal equilibrium , but an ideal gray body, an idealisation which is applicable to heat transfer in cryostats.

The skin depth is the distance an electromagnetic wave propagates into a surface for the flux density to fall by a factor 1/e. In metals the skin depth is given by (see for example Lipson and Lipson 1969) :

$$l = \left(\frac{4\pi^2 c \mu \mu_o \sigma}{\lambda} \right)^{-\frac{1}{2}} \tag{A.12}$$

where μ is the relative permeability ($\sim$1 for non–magnetic materials)

μ_o is the permeability of free space (4πx10^{-7} Hm^{-1} (Ns^2C^{-2}))

σ is the conductivity (5x$10^8 \Omega^{-1}m^{-1}$ for aluminium at 80K)

In metal surfaces (conductors) the skin depth is very small. For 10 μm radiation (λ_{max} for 300K gray body) incident on aluminium at 80K, l is about 0.001 μm. In general for the metal surfaces and radiation fields encountered in cryostats, the transmission $\tau = 0$. Therefore applying Kirchoff's law to Equation A.5 gives :

$$1 - \rho = \alpha = \varepsilon \tag{A.13}$$

From which it follows that highly reflective surfaces have low emissivities. The classical theory of Drude (Kittel 1956) in which the free electrons in a metal are treated as an "electron gas" yields a relation between the reflectivity, wavelength (μm) and electrical conductivity (Ω^{-1} m^{-1}) :

$$\alpha = 1 - \rho = 365(\sigma\lambda)^{-\frac{1}{2}} \tag{A.14}$$

At low temperatures, particularly for infrared and shorter wavelength radiation, the absorbtivity is not as small as this theory predicts because of the anomalous skin effect. This effect occurs when the electron mean free path is long compared with the incident wavelength. However, the conclusions remain valid – metals with high electrical conductivities, such as gold and silver, have the lowest emissivities.

Of basic importance in radiant heat transfer is the *direct* exchange of energy between surfaces – the fraction of the energy leaving one surface which strikes the second surface directly, none being considered which is transferred by reflection or reradiation from other surfaces which may be present. The configuration (or shape) factor F_{2-1} accounts for the emissivities of the surfaces and their geometrical relationship. The geometrical part of the configuration factor is found by determining the exchange between different differential area elements in each surface and then integrating over both surfaces. The geometry of the situation is illustrated in Figure A.2.

Since df is the portion of the flux from dA_2 intercepted by dA_1, it may be written as :

$$df = \frac{dq_{2-1}}{dA_1} = I_2 \frac{\cos\theta_1 \cos\theta_2}{r^2} dA_2 \tag{A.15}$$

Therefore the heat flow is given by :

$$q_{2-1} = \int_{A_1} \int_{A_2} I_2 \frac{\cos\theta_1 \cos\theta_2}{r^2} dA_1 dA_2 \tag{A.16}$$

If the flux originating from surface A_2 (both emitted and reflected),J_2, is uniform over that surface, then the fraction of energy leaving A_2 which is intercepted by A_1 is :

$$F_{2-1} = \frac{q_{2-1}}{A_2 J_2} \tag{A.17}$$

If, in addition, this radiated energy is diffuse, then $J_2 = \pi I_2$. Thus :

$$F_{2-1} = \frac{1}{\pi A_2} \int_{A_1} \int_{A_2} \frac{\cos\theta_1 \cos\theta_2}{r^2} dA_1 dA_2 \tag{A.18}$$

where F_{2-1} is the geometric part of the diffuse configuration factor. Applying Equation A.18 to the infinitely long co–axial cylinder illustrated in Figure A.3, where the polar axis of the spherical coordinate system is aligned parallel to the long axis of the cylinders. O is a point in the differential area dA_2 on the outer cylinder radius R_2, which completely encloses an inner cylinder radius R_1. From spherical coordinates :

$$\cos\alpha_2 = \cos(90 - \theta)\cos\phi$$

and the solid angle subtended at O by the differential area dA_1 at P on the inner cylinder is :

$$\frac{\cos\alpha_1 dA_1}{r^2} = \sin\theta d\theta d\phi$$

where the limits for two infinite co–axial cylinders are $\theta_{1,2} = 0, \pi$ and $\phi_{1,2} = \pm \sin^{-1}\left(\frac{R_1}{R_2}\right)$. Therefore :

$$F_{2-1} = \frac{1}{\pi A_2} \int_0^{A_2} \int_{\phi_1}^{\phi_2} \int_{\theta_1}^{\theta_2} \sin^2\theta \cos\phi d\theta d\phi dA = \frac{R_1}{R_2} = \frac{A_1}{A_2} \tag{A.19}$$

Similarly for concentric spheres $F_{2-1} = \left(\frac{R_1}{R_2}\right)^2 = \frac{A_1}{A_2}$. Note that in neither case can the inner surface 'see' itself ($F_{2-2} = 0$). Of the radiation $\sigma A_2 \varepsilon_2 T_2^4$ emitted by the outer body, the inner body absorbs a fraction $(A_1/A_2)\varepsilon_1$ and reflects a fraction $(1 - A_1\varepsilon_1/A_2)$ to the outer body which intercepts all of it and reflects a fraction $(1 - \varepsilon_2)$ of which a further fraction $(A_1/A_2)\varepsilon_1$ is absorbed by the inner body, and so on. The total

radiation received by the inner body is therefore the sum of the series :

$$\sigma A_2\varepsilon_2 T_2^4\left(\frac{A_1}{A_2}\right)\varepsilon_1+\sigma A_2\varepsilon_2 T_2^4\left(1-\frac{A_1}{A_2}\varepsilon_1\right)(1-\varepsilon_2)\left(\frac{A_1}{A_2}\right)\varepsilon_1$$
$$+\sigma A_2\varepsilon_2 T_2^4\left(1-\frac{A_1}{A_2}\varepsilon_1\right)^2(1-\varepsilon_2)^2\left(\frac{A_1}{A_2}\right)\varepsilon_1+\ldots \qquad \text{(A.20)}$$

This is a geometric progression with a common ratio :

$$r=\left(1-\frac{A_1}{A_2}\varepsilon_1\right)(1-\varepsilon_2) \text{ and the sum to n terms } S_n=a\frac{r^n-1}{r-1}$$

where a is the first term in the series and r is less than unity. Therefore the sum to infinity and the net transfer to the inner body is :

$$q=\sigma A_1\left[\frac{1}{1/\varepsilon_1+(A_1/A_2)(1/\varepsilon_2-1)}\right](T_2^4-T_1^4) \qquad \text{(A.21)}$$

Thus :

$$F_{2-1}=\left[\frac{1}{1/\varepsilon_1+(A_1/A_2)(1/\varepsilon_2-1)}\right] \qquad \text{(A.22)}$$

When $A_2 \gg A_1$, F_{2-1} reduces to ε_1, and in such a case the heat load radiated onto the inner body is independent of the nature of the surface of the outer body (the cavity approximation). The affect of introducing a number (n) of "floating' radiation shields between the inner and outer bodies can be investigated by rewriting Equation A.21 for the i^{th} in the form :

$$\frac{q\left[1/\varepsilon_{i+1}+(A_{i+1}/A_{i+2})(1/\varepsilon_{i+2}-1)\right]}{\sigma A_{i+1}}=T_{i+2}^4-T_{i+1}^4 \qquad \text{(A.23)}$$

This is an idealisation in which the only heat transfer is by radiation. When thermal equilibrium is reached the heat flow q is constant through all the separate shields. Summing over the n shields yields :

$$\frac{q}{\sigma}\sum_{i=1}^{i=n+1}\frac{1}{A_i}\left[\frac{1}{\varepsilon_i}+\frac{A_i}{A_{i+1}}\left(\frac{1}{\varepsilon_{i+1}}-1\right)\right]=T_{n+2}^4-T_1^4 \qquad \text{(A.24)}$$

If the emissivities are all the same and $A_1=A_{n+1}$ then :

$$\frac{q}{A_1}=\frac{1}{n+1}\sigma\left(\frac{1}{2/\varepsilon-1}\right)(T_{n+2}^4-T_1^4) \qquad \text{(A.25)}$$

References

Adkins, C.J., 1975. *Equilibrium Thermodymamics*,
McGraw–Hill Book Company (UK) Limited.

Barron, R.F., 1985. *Cryogenic Systems*, Oxford University Press, New York.

Batrla, W., Wilson, T.L., Batien, P. & Ruf, K., 1983. *Astr. Astrophys.*, **128**, 128.

Boyd, R.W., 1983. *Radiometry and the Detection of Optical Radiation*,
John Wiley & Sons, New York.

Castelaz, M.W., Hackwell, J.A., Grasdalen, G.L., Gehrz, R.D. & Gullixson, C.,
1985. *Astrophys J.*, **290**, 261.

Castelaz, M.W., Hackwell, J.A., Grasdalen, G.L. & Gehrz, R.D., 1986. *Astrophys J.*,
300, 406.

Chapman, A.J., 1984. *Heat Transfer*, Macmillan Publishing Company, New York.

Clausing, A.M. & Chao, B.T., 1965. *J. of Heat Transfer.*, **27**, 243.

Cohen, J.G. & Frogel, J.A., 1977. *Astrophys. J.*, **211**, 178.

Cohen, M. & Kuhi, L.V., 1979 . *Astrophys. J.*, **41**, 743.

Corruccini, R.J., 1959. *Vacuum*, **7–8**, 19.

CTI–Cryotronics, 1984. *Installation, operation and servicing instructions:
Cryocooler 350C.*

Daunt, J.G. & Rosen, C.Z., 1970. *J. of Low Temperature Physics.* **3**, 89.

Davis, L. & Greenstein, J.L., 1951. *Astrophys. J.*, **114**, 206.

Dyck, H.M. & Lonsdale, C.J., 1979. *Astr. J.*, **84**, 1339.

Fowler, A.M., Britt, J.P., Joyce, R.R., Probst, R.G. & Gates G.L., 1986.
In: *Proc. Second Infrared Detector Technology*, NASA (88213), 25–1.

Fowler, A.M., Probst, R.G.Britt, J.P., Joyce, R.R., & Gillett F.C., 1987.
In *Optical Engineering.* **26**, 232.

Fischer, J., Righini–Cohen, G. & Simon, M., 1980. *Astrophys. J.*, **238**, L155.

Fischer, J., Sanders, D.B., Simon, M. & Solomon, P.M., 1985. *Astrophys. J.*, **293**, 508.

Forrest, W.J. & Shure, M.A, 1986. *Astrophys. J.*, **311**, L81.

Gatley, I., Becklin E.E., Matthews, K., Neugebauer, G., Penston, M.V.,
& Scoville, N., 1974. *Astrophys. J.*, **191**, L121.

Genzel, R. & Downes, D., 1979. *Astr. Astrophys.*, **72**, 234.

Goodall, D.H.J., 1970. *Cryogenic Data Chart*, Culham Laboratory.

Goudis, C., 1982. *The Orion Complex: A case study of interstelllar matter.* D. Reidel Publishing Company, Dordrecht, Holland.

Hilberath, W. & Vowinkel, B., 1983. *Cryogenics*, **23**, 467.

Hough, J.H., Axon, D.J., Burton, M.G., Gatley, I., Sato, S., Bailey, J., McCaughrean, M.J., McLean, I.S., Nagata, T., Allen, D., Garden, R.P., Hasegawa, T., Hayashi, M., Kaifu, N., Morimoto, M. & Walther, D., 1986. *Mon. Not. R. astr. Soc.*, **229**, 526.

Hough, J.H., Sato, S., Tamura, M., Yamashita, T., McFadzean, A.D., Rouse, M.F., Whittet, D.C., Kaifu, N., Suzuki, H., Nagata, T., Gatley, I. & Bailey. J., 1988. *Mon. Not. R. astr. Soc.*, **230**, 107.

Hubble, E.P., 1922. *Astrophys. J.*, **56**, 400.

Kittel, C., 1956. *Intro. to Solid State Physics*, John Wiley & Sons Inc., New York.

Kornneef, J., 1983. *Astr. Astrophys.*, **128**, 84.

Kutner, M.L., Evans II, N.J. & Tucker, K.D., 1986. *Astrophys. J.*, **288**, 456.

Lamb, G., Shu, P., Mather J., Ewin, A. & Bowser, J., 1987. In: *Proc. Infrared Astronomy with Arrays*, University of Hawaii.

Lipson, S.G. & Lipson, H., 1969. *Optical Physics*, Cambridge University Press.

McCaughrean, M.J., 1988. *Ph. D. Thesis*, University of Edinburgh.

McDonell, M.W. & Klee, H.W., 1984. *Optical Engineering*, **23**, 187.

McLean, I.S., Chuter, T.C, McCaughrean, M.J. & Rayner, J.T., 1986. In: *Instrumentation in Astronomy VI, SPIE*, **627**, 430.

McLean, I.S., 1987. In: *Proc. Infrared Astronomy with Arrays*, University of Hawaii.

Meaburn, J., 1976. *Detection and Spectroscopy of Faint Light.* D. Reidel Publishing Company, Dordrecht, Holland.

Mestel, L. & Paris, R.B., 1984. *Astr. Astrophys.*, **136**, 98.

Mestel, L. & Ray, T.P., 1985. *Mon. Not. R. astr. Soc.* **212** 275.

Orias, G., Hoffman, A. W. & Casselman, M.F, 1986. In: *Instrumentation in Astronomy VI, SPIE*, **627**, 408.

Nast, T., Bell, G. & Barnes, C., 1982. *Advances in Cryogenic Engineering*, **27**, 1117.

Parenago, P.P., 1954. In *Trudy Gosudarstvennogo Astronomicheskogo Instituta im. P.K. Shternberga, Moscow Universitet*, **25**.

Pendleton, Y., Werner, M.W., Capps, R. & Lester, D., 1986. *Astrophys, J.* **311**, 360.

Purcell, E.M., 1975. In: *The Dusty Universe*, eds. Cameron,G.B. & Cameron, A.G.W., Neal Watson, New York.

Purcell, E.M., 1979. *Astrophys. J.*, **231**, 404.

Rayner, J.T., 1984. *M. Sc. Dissertation*, University of Edinburgh.

Rayner, J.T., Chuter, T.C., McLean, I.S., Radostitz, J.V., & Nolt, I.G., 1988. *Infrared Physics*, **28**, 43.

Rieke, G.H. & Lebofsky, M.J., 1985. *Astrophys. J.*, **288**, 618.

Roark, R., 1956. *Formulas in Stress and Strain*, McGraw–Hill Book Company, New York.

Roberts, J.K. & Miller, A.R., 1960. *Heat and Thermodynamics*, Blackie & Son Limited, Glasgow.

Rode, J.P., Blackwell, J.D., Blessinger, M.A. & Vural, K., 1987. In: *Proc. Infrared Astronomy with Arrays*, University of Hawaii.

Rohsenow, W.M. & Hartnett, J.D. (eds), 1973. *Handbook of Heat Transfer*, McGraw–Hill Book Company,p. 3–14ff.

Rose–Innes, A.C., 1973. *Low Temperature Laboratory Techniques*, English University Press, London, p. 155.

Scott, R.B., 1959. *Cryogenic Engineering*, D. Van Nostrand Co., Inc., Princeton, N.J.

Shu, Q.S., Fast, R.W. & Hart, H.L., 1986. *Cryogenics*, **26**, 671.

Shu, Q.S., 1987. *Cryogenics*, **27**, 249.

Shu, Q.S., Fast, R.W. & Hart, H.L., 1987. *Cryogenics*, **27**, 298.

Solomon, P.M., Huguenin, G.R. & Scoville, N.Z., 1981. *Astrophys. J.*, **245**, L19.

Smith, J., Lynch, D.K., Cudaback, D. & Werner, M.W., 1979. *Astrophys. J.*, **234** 902.

Thom, R.D. & Yang, B.T., 1986. In: *Proc. Second Infrared Detector Technology.* NASA (88213), 25–1.

Torre, J.P. & Chanin G., 1984. Rev. Sci. Instr., **55**, 213.

Thronson, H.A., Harper, D.A., Keene, J., Loewenstein, R.F., Mosely, H. & Telesco, C.M., 1978. *Astr. J.*, **83**, 492.

Thronson, H.A. & Thompson, R.I., 1982. *Astrophys. J.*, **254**, 543.

Vampola, J.L., 1982. *InSb Detector Model for ADS*, Internal Memorandum, SBRC.

Ward–Thompson, D., Warren–Smith, R.F., Scarrot, S.M. & Wolstencroft, R.D., 1985. *Mon. Not. R. astr. Soc.*, **215**, 98.

Walker, G.K, 1983. *Cryocoolers*, Plenum Press, New York.

Westbrook, W.E., Werner, M.W., Elias, J.H., Gezari, D.Y., Hauser, M.G., Lo, M.G. & Neugebauer, G., 1976. *Astrophys. J.*,**209** 94.

Wilner, S.P., Gillett, F.C., Herter, T.L., Jones, B., Krassner, J., Merrill, K.M., Pipher, J.L., Puetter, R.C., Rudy, R.J., Russell, R.W & Soifer, B.T. *Astrophys. J.*, **253** 174.

White, G.K., 1979. *Experimental Techniques in Low Temperature Physics*, Clarendon Press, Oxford.

Wolfe, W.L. & Zissis, G.J. (eds), 1978. The Infrared Handbook, Office of Naval Research, Dept. of the Navy, Washington D. C.

Yang, E.S., 1978. *Fundamentals of Semiconductor Physics*, McGraw–Hill.